The Dissertation Committee for Andrew Adam Roth Certifies that this is the approved version of the following dissertation:

Snu40p and Snu66p are Required for Spliceosome Activation at Suboptimal Temperatures

Committee:

Scott W. Stevens, Supervisor

Dean Appling

Walter Fast

Edward Marcotte

Whitney Yin

Snu40p and Snu66p are Required for Spliceosome Activation at Suboptimal Temperatures

by

Andrew Adam Roth, B.S.; M.B.A.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

May 2008

Dedication

For Suzie

Acknowledgements

I would like to express my gratitude to Scott Stevens for accepting me into his laboratory and seeing me through the trials and tribulations inherent to the experience. I would also like to thank the other members of the Stevens lab past and present for expertise, encouragement, and entertainment over the years. Individually, they are Grace Chen, Josh Combs, Champ Gupton, Jennifer Hennigan and Rea Lardelli.

I would also like to thank Samantha Croft and Jeff Croarkin in the Gottlieb lab for sharing with me casual conversation and lab space over the years. The advice, friendship, and lively debate with Patrick Killion made the journey an enjoyable one.

The Yates lab at the Scripps research institute deserves special recognition for their effort in analyzing my protein samples in a capable and timely manner. Truly, this work would not have been complete without their efforts.

I thank the members of my dissertation committee for their time and guidance: Dean Appling, Walter Fast, Edward Marcotte, and Whitney Yin.

The input of Arlen Johnson and the members of his lab, past and present, have been invaluable. Among them: John Hedges, Matt West, Alice Wang, Ivy Hung, Kai-Yin Lo, Cyril Bussiere, and Josh White. The Lambowitz, Iyer, Krug, Marcotte, Paull, and Bose labs all deserve thanks for their impromptu help on technical and resource matters.

I would like to thank Cecil Harkey for unfailing support and generally running a facility worthy of the term ‘excellent’. Lastly, I would like to thank Bill Cassady for his professional advice and diligence in and around campus.

Snu40p and Snu66p are Required for Spliceosome Activation at Suboptimal Temperatures

Publication No._____

Andrew Adam Roth, Ph.D.

The University of Texas at Austin, 2008

Supervisor: Scott W. Stevens

In addressing the pre-mRNA substrate, the splicing machinery requires rearrangement of multiple RNA and protein components. The classical model of spliceosome formation begins with the U1 snRNA recognition of the 5' splice site and U2 snRNP interaction with the branch point. This process is followed by the engagement of a pre-assembled U4/U6•U5 tri-snRNP to form the A2-1 complex. The spliceosome is subsequently activated through a number of structural rearrangements. Among these is the unwinding of the U4/U6 intermolecular helix by the tri-snRNP component Brr2p.

While numerous protein components of the tri-snRNP have been identified, the function of many of these remain unknown. The nonessential Snu66p (U4/U6•U5-110K in humans) stably associates only with the U4/U6•U5 tri-snRNP while the similarly nonessential Snu40p (U5-52K in humans) associates exclusively with the U5 snRNP. To understand why two non-essential pre-mRNA splicing factors have been so well conserved through great evolutionary distances, we examined their roles in the assembly

and function of the tri-snRNP. Removal of *SNU40* alone does not affect snRNP levels, however deletion of *SNU66* results in reduced levels of tri-snRNP. The U4/U6•U5 snRNPs in *Δsnu66* cells are resistant to the ATP-dependent U4/U6 unwinding by Brr2p, and profound U4/U6 accumulation occurs at reduced temperatures. Remarkably, subsequent removal of *SNU40* in a *Δsnu66* strain bypasses the tri-snRNP formation defect while unwinding of U4/U6 remains defective. Additional investigation revealed that Prp6p, another tri-snRNP protein, is destabilized from the complex. Based upon this data in total, I present a model in which Snu40p and Snu66p interact sequentially with Prp6p to maintain directionality for proper biogenesis of the tri-snRNP.

Further, the U4/U6 unwinding defect of the double mutant should theoretically arrest the A2-1 spliceosome. Indeed, native gel analysis confirms the buildup of a large complex later determined to be A2-1. I have purified this complex, functionally tested its catalytic viability, and identified its components via mass spectrometry. This is the first full characterization of the A2-1 precatalytic spliceosome complex in *Saccharomyces cerevisiae*.

Table of Contents

List of Tables

List of Figures

List of Illustrations

Chapter 1: General Introduction

1.1 Overview of Splicing

1.1.1 Introduction

It is estimated that up to 30% of human disease causing point mutations affect the process of pre-mRNA splicing (Krawczak, Reiss, and Cooper 1992)(Nissim-Rafinia and Kerem 2005). While this fact alone would make the process worthy of investigation, the field presents potential answers to questions beyond utilitarian disease treatment.

Although not a defining characteristic of the eukaryotic domain (Yoshinari, Takashi Itoh, Hallam, DeLong, Yokobori, Yamagishi, Oshima, Kita, and Yoh-ichi Watanabe 2006), the frequency and complexity of splicing seems to correlate with evolutionary complexity (Eddo Kim, Magen, and Ast 2007). The genes of many single cell eukaryotes often carry a miniscule number of introns while in mammals splicing is almost exclusively a requirement. Even beyond the simple intron, however, it appears that alternative splicing follows this same trend. There are no known proteins that result from an alternatively spliced mRNA in the single celled brewer's yeast *Saccharomyces cerevisiae.* In contrast, 2% of proteins result from alternatively spliced mRNA in the worm and 7% in the fly (Harrison, Kumar, Lang, Snyder, and Gerstein 2002) It is estimated that greater than 50% of all human genes are alternatively spliced (Phillip A. Sharp 2005). On the basis of these simple correlations, it would appear that the process of mRNA splicing has wide implication in phylogenetics and organism complexity.

Nevertheless, it is the complexity of the process itself that has has made the task of understanding it so difficult. Indeed, over thirty years of research has yielded only a limited understanding of the process.

1.1.2 History

In 1977 the work of Phillip Sharp (Berget, C Moore, and P A Sharp 1977), Richard Roberts (L T Chow, Gelinas, Broker, and R J Roberts 1977) and colleagues first described what was then known as the 'split gene'. Their work compared DNA sequences and the mRNA that ultimately resulted from them. It was found that large intervening sequences were removed prior to the formation of mature mRNA. These removed RNA sequences were later termed 'introns' by Wally Gilbert (Phillip A. Sharp 2005). Similarly, the RNA sections remaining after splicing were termed 'exons'. Soon thereafter, the importance of mRNA splicing in human biology was underscored when a number of β-thalassemia causing defects were found to be the result of splicing defects (Treisman, Orkin, and Maniatis 1983). By the time the existence of the split gene was uncovered in the late 1970's, the splicing machinery itself had been under investigation for over a decade. It wasn't until the 1980's, however, that it became clear that snRNAs and snRNPs were the instrument of mRNA splicing (Lerner, Boyle, Mount, Wolin, and J A Steitz 1980).

1.1.2.1 RNA Component Discovery

In 1965, Okamura and Busch published an analysis of RNA base composition in tissue from Walker 256 rats (Okamura and H Busch 1965). This work noted that whole nuclear extract was rich in uridylic acid. This data was quickly followed by a second analysis (Muramatsu, James L. Hodnett, and Harris Busch 1966) showing that these U-RNAs sedimented in a sucrose gradient between 4S-8S. The first of the snRNAs to be

isolated were U1, U2, and U3 in 1968 (J L Hodnett and H Busch 1968). During this isolation, it was noted that while U1 and U2 were found in the nucleoplasm, U3 was restricted to the nucleolus. By the mid-1970's, snRNAs (Small Nuclear RNAs) were identified and localized (Zieve and Penman 1976) within the HeLa cell. Finally, by late 1970s and early 1980s most of the snRNAs had been sequenced (U1 (C Branlant, A Krol, Ebel, E Lazar, H Gallinaro, M Jacob, Sri-Widada, and Jeanteur 1980),U2 (Shibata, T. S. Ro-Choi, R. Reddy, Y. C. Choi, Henning, and H. Busch 1975), U3 (R Reddy, Henning, and H Busch 1979), U4 (R Reddy, Henning, and H Busch 1981), U5 (Alain Krol, Helene Gallinaro, Eliane Lazar, Monique Jacob, and Christiane Branlant 1981), U6 (Epstein, R Reddy, Henning, and H Busch 1980)).

In 1980 the laboratory of Joan Steitz published work (Lerner, Boyle, Mount, Wolin, and J A Steitz 1980) comparing the consensus splice junction sequences with the 5' sequence of the U1 snRNAs. The complementarity naturally led to speculation that this RNA species played a role in the splicing process. Also included in this work was evidence that U1, U2, U4, U5, and U6 snRNPs were all bound by anti-Sm antibodies as well as antibodies from patients with lupus erythematosus. This latter data indicated that all five of these snRNAs were involved in the same process. Additionally it was noted that U3 was not included in this list. Earlier analysis of U3 demonstrated its ability to base pare with the 28S pre-rRNA (Prestayko, Tonato, and H Busch 1970). Thus, it was inferred that U3 was not likely to be involved in mRNA splicing but the rRNA maturation process (Riedel, Wise, Swerdlow, Mak, and C Guthrie 1986).

Later work discovered that the process of intron removal utilized a lariat RNA intermediate with a corresponding 5' exon (Padgett, Grabowski, M M Konarska, Seiler, and P A Sharp 1986).

1.1.2.2 Protein Component Discovery

After the discovery of U-RNAs in the mid 1960s, work commenced on the identification of interacting partners (H Busch, R Reddy, Rothblum, and Y C Choi 1982). It was rapidly recognized that these RNAs were complexed with a number of protein components when it was found that protease treatment reduced the RNA sedimentation rate (Enger and Walters 1970). These proteins were not isolated, however, until a 1975 paper on the isolation of the U1 and U2 snRNPs (Raj, Tae Suk Ro-Choi, and Harris Busch 1975). The biggest advance to date in snRNP biochemistry came in 1979 with the realization that antibodies from human lupus erythematosus patients recognized snRNP components (Lerner and J.A. Steitz 1979). This allowed the efficient purification of snRNPs and their component snRNAs (Lerner, Boyle, Hardin, and J.A. Steitz 1980).

1.1.3 General Concepts in Splicing

1.1.3.1 DEAD/H Box Helicases

The two transesterification reactions of mRNA splicing do not technically require the input of energy. Nevertheless, ATP is required throughout the splicing cycle by the helicases responsible for the RNA rearrangements that make splicing possible. These helicases appear to be evolutionarily related and share what is known as the DEAD/H box domain (Patrick Linder 2006). Four of the proteins share the DEAD box motif: Sub2p, Prp5p, Prp28p, Brr2p. (West and Milgrom 2002)(de la Cruz, Kressler, and P. Linder 1999)(Patrick Linder 2006). The related DEAH box domain is found in the remaining three helicases: Prp2p, Prp16p, and Prp22p (Krainer 1997). While the larger class of DExH/D helicases are best known for the ability to unwind RNA-RNA duplexes, they also have the ability to perturb RNA-protein interactions (Jankowsky, Gross, Shuman, and Pyle 2001).

1.1.3.2 snRNP Biogenesis

The U1, U2, U4, and U5 snRNAs are transcribed by RNA polymerase II, while the U6 snRNA is transcribed by RNA Polymerase III. Likely a result of this disparity in transcription, the overall biogenesis pathway also differs for U1, U2, U4, and U5 versus the Pol III transcribed U6 (Tamas Kiss 2004).

The snRNAs that result from RNA Polymerase II are immediately capped with a monomethyl guanosine cap (m^7G). This cap attaches to the Cap Binding Complex (CBP), which acts as an adapter for the combined PHAX (Phosphorylated Adapter for RNA Export) and CRM1/RanGTP export machinery (Ohno, Segref, Bachi, Wilm, and I W Mattaj 2000). This complex then shuttles the nascent snRNA from the nucleus via the nuclear pore complex. Once in the cytoplasm, the snRNA export complex disassembles and the snRNA is addressed by the Survival of Motor Neurons (SMN) complex (Massenet, Pellizzoni, Paushkin, Iain W Mattaj, and Dreyfuss 2002)(Paushkin, Gubitz, Massenet, and Dreyfuss 2002). This machinery assists the loading of the Sm complex protein components onto the Sm binding site of the snRNA. By nature of their common biogenesis process, U1, U2, U4, and U5 all share the same Sm protein ring components. This group of seven proteins includes Smb1p, Smd1p, Smd2p, Smd3p, Sme1p, SmFp, and SmGp which bind to the conserved Sm motif on the snRNAs. The binding of the Sm core is a precondition for the activity of the Tgs1p methyltransferase and 3' end processing (Tamas Kiss 2004). Tgs1p hypermethylates the 5' cap to a 2,2,7-trimethylguanosine (TMG) while it is believed that RNAse III is ultimately responsible for cleavage of the 3' end to its final state (Chanfreau, S A Elela, Ares, and C Guthrie 1997). After this post-processing is complete, Snuportin1 works through a importin-β dependant mechanism (Palacios, Hetzer, Adam, and I W Mattaj 1997) to the re-import of the snRNP into the nucleus (Matera and Shpargel 2006).

The biogenesis of the RNA Polymerase III produced U6 snRNA differs from the other Pol II produced transcripts, most notably as it is not exported from the nucleus as part of its biogenesis. Immediately after transcription U6 is bound by the La protein, which is then displaced by the formation of the Lsm protein ring (made of Lsm2p, Lsm3p, Lsm4p, Lsm5p, Lsm6p, Lsm7p, Lsm8p) near the 3' end of U6. The formation of this heptameric ring of Lsm proteins is responsible for U6 transport to the nucleolus where the snRNA is methylated (B E Jády and T Kiss 2000) and pseudouridylated (Arnold M Kiss, Beáta E Jády, Bertrand, and Tamás Kiss 2004).

All five snRNAs are adorned with a changing array of protein partners. Some proteins have been reported to be exclusive to a single particle, while others remain paired throughout the splicing cycle. Table 1.1 lists the proteins that have been identified as members of the U1, U2, U4/U6, U5, U6, and tri-snRNP complexes. U4 has not been characterized separately from its union with the U6 snRNP.

U1 Proteins
Luc7p
Mud1p
Nam8p
Prp39p
Prp40p
Prp42p
Smb1p
Smd1p
Smd2p
Smd3p
Sme1p
SmFp
SmGp
Snp1
Snu56p
Snu71p
Yhc1p

U2 Proteins
Cus1p
Cus2p
Hsh155p
Hsh49p
Ist3p
Lea1p
Msl1p
Prp11p
Prp21p
Prp9p
Smb1p
Smd1p
Smd2p
Smd3p
Sme1p
SmFp
SmGp
Rse1p
Snu17p
Ysf3p

U4-U6 Proteins
Lsm2p
Lsm3p
Lsm4p
Lsm5p
Lsm6p
Lsm7p
Lsm8p
Prp3p
Prp31p
Prp4p
Smb1p
Smd1p
Smd2p
Smd3p
Sme1p
SmFp
SmGp
Snu13p

U5 Proteins
Prp8p
Brr2p
Snu114p
Prp28p
Snu40p
Dib1p
Aar2p
Prp6p*
Prp18p
Smb1p
Smd1p
Smd2p
Smd3p
Sme1p
SmFp
SmGp

U6 Proteins
Lsm2
Lsm3
Lsm4
Lsm5
Lsm6
Lsm7
Lsm8
Prp24p

tri-snRNP Proteins
Brr2p
Dib1p
Lsm2p
Lsm3p
Lsm4p
Lsm5p
Lsm6p
Lsm7p
Lsm8p
Prp3p
Prp31p
Prp38p
Prp4p
Prp6p
Prp8p
Sad1p
Smb1p
Smd1p
Smd2p
Smd3p
Sme1p
SmFp
SmGp
Snu114p
Snu13p
Snu23p
Snu66p
Spp381p

Table 1.1 - Core snRNP proteins

Data based on (Eurie L Hong et al. 2008)(Krainer 1997)(Alexander Gottschalk, Bartels, Gitte Neubauer, Reinhard Luhrmann, and Patrizia Fabrizio 2001)(S.B. Liu, Rauhut, H.P. Vornlocher, and R. Lührmann 2006). *Although the localization of yeast Prp6p has been a matter of debate, our work in Chapter 3 indicates that it transiently interacts with U5 to the exclusion of the U4/U6 snRNP.

1.1.4 The Splicing Cycle

1.1.4.1 Overview

The two transesterification reactions of the mRNA splicing reaction are driven by the formation and remodeling of the spliceosome. As the components are utilized and remodeled for future activity, the events form a cycle. The mRNA splicing cycle has historically been represented as a stepwise progression of players joining and later exiting the substrate mRNA (R.J. Lin, Newman, S.C. Cheng, and J. Abelson 1985). More recent data raises the possibility that the spliceosome is formed prior to addressing the substrate (Scott W. Stevens, Ryan, Helen Y. Ge, Roger E. Moore, Mary K. Young, Terry D. Lee, and John Abelson 2002). In this scenario, the formation and recycling events take place as a series of rearrangements rather than physical union and separation. As the newer preformed model is addressed later in this work in section 1.1.3.6, this section will review the cycle as it is known under the traditional stepwise model.

During transcription the carboxy-terminal domain (CTD) of RNA polymerase II recruits enzymes that create a 7-metyl guanosine cap on the 5' end of the nascent pre-mRNA (Cho, Takagi, Christine R. Moore, and Buratowski 1997)(Fong and Bentley 2001). After this capping but while transcription continues it is believed that the CTD recruits the protein components required for splicing (Listerman, Sapra, and K.M. Neugebauer 2006).

The U1 snRNP binds the 5' splice site thus forming the commitment complex. The pre-spliceosome or B complex is formed with the ATP dependent addition of the U2 snRNP (S.C. Cheng and J. Abelson 1987). The addition of U2 is enabled by the splicing factor U2AF, which recognizes the polypyrimidine tract between the branchpoint and the

3’ splice site. The addition of a preformed U4/U6·U5 tri-snRNP completes the A2-1 complex. A series of ATP and GTP dependent rearrangements result in the unwinding of U4 from U6, followed by the release of U1 and U4 from the complex. This frees the 5’ splice site to be addressed by the U6 snRNA which also establishes base pairing with U2 (D.A. Wassarman and J.A. Steitz 1993). The complex is now referred to as the activated spliceosome or A1 complex. The helicase Prp2p causes a rearrangement that leads to the A2-2 complex in preparation for the first step of splicing (Silverman, Maeda, Wei, P. Smith, J.D. Beggs, and R.J. Lin 2004). The first transesterification reaction leads to the A2-3 complex, which contains the lariat intermediate and a free 5’ exon. The second step of splicing then completes the reaction and the component snRNPs are released for recycling.

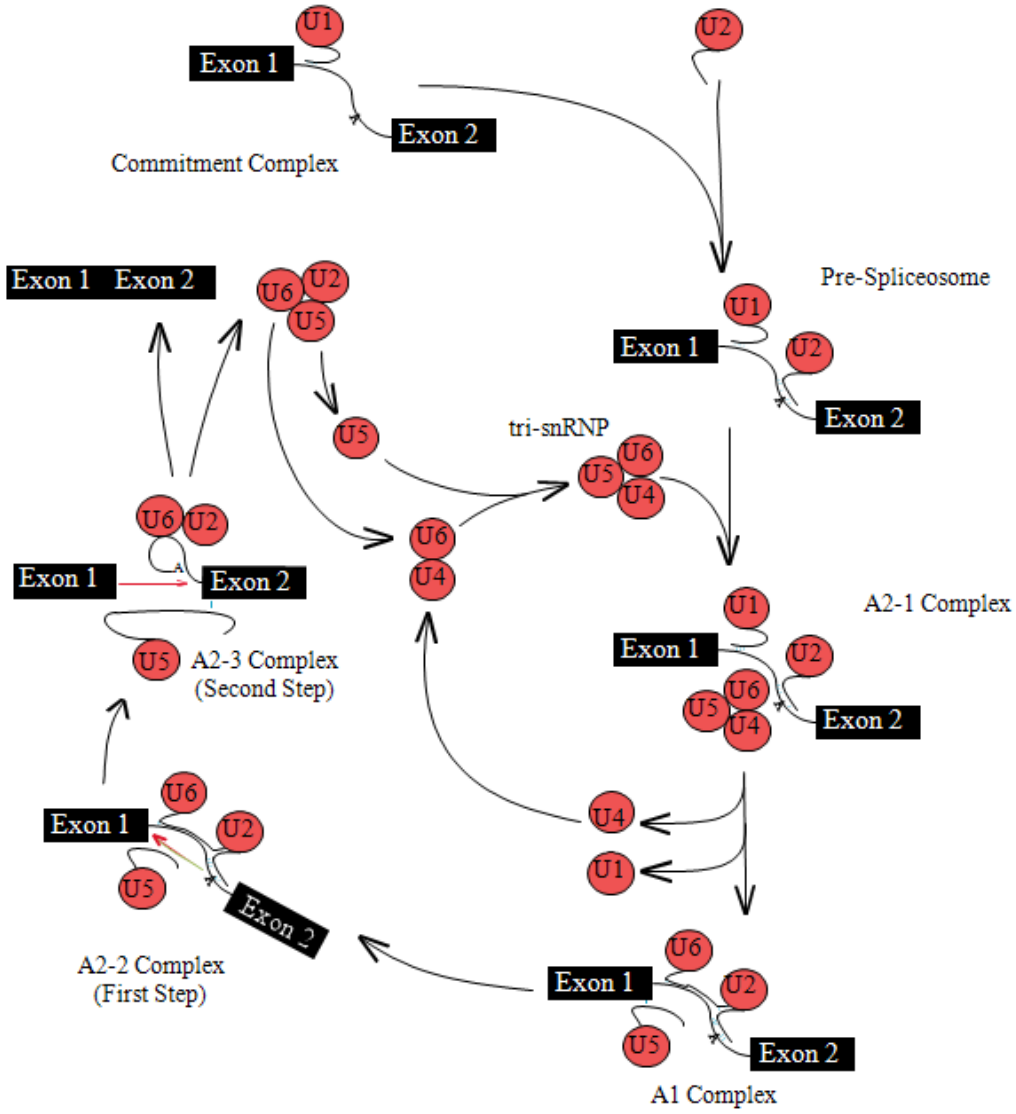


Illustration 1.1 - The mRNA splicing cycle

Adapted from Scott Stevens

1.1.4.2 The U4/U6 snRNP

After splicing is completed, the U6 snRNA is recycled where it joins newly synthesized U6 in being annealed to the U4 snRNA. The process of reannealing U4 to U6 requires the Prp24 protein (Rader and Christine Guthrie 2002). The end result of the U4/U6 annealing process is a 13S particle containing not only the two snRNAs but the proteins Prp3p, Prp4p, Prp31, and Snu13p in addition to the Sm and Lsm core proteins (Table 1.1). While the RNA base pairing at the core of this particle no doubt provides a good deal of structural integrity, a protein-protein interaction between the human homologs of Prp4p and Prp3p have been mapped by co-immunoprecipitation as well (Gonzalez-Santos, Anan Wang, Joses Jones, Ushida, Jun Liu, and Hu 2002).

1.1.4.3 The U4/U6·U5 tri-snRNP

Unlike the basepaired U4/U6 di-snRNP, the union of U4/U6 with the 16S U5 snRNP to create the 25S tri-snRNP does not rely on basepaired interactions. Rather, a series of protein-protein interactions is thought to bind the particles together. While the structures of U4/U6 and the tri-snRNP remain unsolved, some of the protein-protein interactions have been mapped. The most prominent of these interactions is amongst the core of the U5 snRNP. The human homologs of Brr2p, Prp8p, and Snu114p form a snRNA-free protein complex that has been purified under high salt conditions (T. Achsel, Ahrens, Brahms, Teigelkamp, and R. Lührmann 1998). In the years after this discovery a number of additional protein-protein interactions have been exposed by co-immunoprecipitation and two-hybrid experiments. The most recent of these datasets, published after the start of this dissertation work, revealed ten additional interactions (S.B. Liu, Rauhut, H.P. Vornlocher, and R. Lührmann 2006). The sum of all the data in this area is presented in Table 1.2. Although these interactions lack functional

verification, some inferences can be drawn. Prp6p and Snu66p are the only identified components to have connections with members of the U5 snRNP and U4/U6 snRNP. This implies that these two tri-snRNP components might bridge members of the former U4/U6 and U5 snRNP components. This hypothesis is both bolstered and made more difficult by the statics and dynamics of these proteins. It is widely accepted that Snu66p and homologs reside exclusively in the tri-snRNP particle. The mammalian Prp6p counterpart, 102K, has been confirmed to be a member of both the U5 and tri-snRNPs. Nevertheless, the yeast Prp6p has been reported to be a member of the tri-snRNP and U4/U6 snRNP to the exclusion of U5. At the beginning of work for this dissertation the reason for this discrepancy was unknown.

Protein 1	Protein 2	Method	Species	Reference
Brr2p	Prp8p	co-ip	human	Achsel 1998
Brr2p	Snu114p	co-ip	human	Achsel 1998
Sn114p	Prp8p	co-ip	human	Achsel 1998
Brr2p	Prp8p	two hybrid	yeast	van Nues 2001
Brr2p	Prp16p	two hybrid	yeast	van Nues 2001
Brr2p	Slu7p	two hybrid	yeast	van Nues 2001
Brr2p	Snp1p	two hybrid	yeast	van Nues 2001
Brr2p	Snu66p	two hybrid	yeast	van Nues 2001
Prp4p	Prp3p	co-ip	human	Gonzales-Santos 2002
Prp31p	Prp6p	co-ip	human	Makarova 2002
Prp31p	Prp6p	two hybrid	human	Makarova 2002
Snu40p	Prp6p	two hybrid	human	Laggerbauer 2005
Snu40p	Dib1	two hybrid	human	Laggerbauer 2005
Snu40p	Prp6p	co-ip	human	Laggerbauer 2005
Snu40p	Dib1	co-ip	human	Laggerbauer 2005
Brr2p	Prp8p	two hybrid	human	Liu 2006
Brr2p	Snu114p	two hybrid	human	Liu 2006
Sn114p	Prp8p	two hybrid	human	Liu 2006
Prp6p	Prp8p	co-ip	human	Liu 2006
Prp6p	Prp8p	two hybrid	human	Liu 2006
Prp6p	Dib1	co-ip	human	Liu 2006
Prp6p	Dib1	two hybrid	human	Liu 2006
Prp6p	Prp3p	co-ip	human	Liu 2006
Prp6p	Snu66p	two hybrid	human	Liu 2006
Prp6p	Snu66p	co-ip	human	Liu 2006
Prp6p	Brr2p	two hybrid	human	Liu 2006
Prp6p	Brr2p	co-ip	human	Liu 2006
Prp6p	Snu114p	two hybrid	human	Liu 2006
Snu66	Prp3p	two hybrid	human	Liu 2006
Snu66	Prp3p	co-ip	human	Liu 2006
Brr2p	Snu66p	co-ip	human	Liu 2006

Table 1.2 – Interactions of the tri-snRNP proteins

Table uses yeast names for simplicity of comparison. Data from (T. Achsel, Ahrens, Brahms, Teigelkamp, and R. Lührmann 1998)(van Nues and Jean D. Beggs 2001)(Gonzalez-Santos et al. 2002)(Olga V Makarova, Evgeny M Makarov, Sunbin Liu, Hans-Peter Vornlocher, and Reinhard Lührmann 2002)(Laggerbauer, Sunbun Liu, Evgeny Makarov, Hans-Peter Vornlocher, Olga Makarova, Ingelfinger, Tilmann Achsel, and Reinhard Luhrmann 2005)(S.B. Liu, Rauhut, H.P. Vornlocher, and R. Lührmann 2006).

1.1.4.4 The A2-1 Complex and Spliceosome Activation

The A2-1 complex is defined by the presence of all five snRNAs including a basepaired U4/U6 in the presence of the substrate pre-mRNA. Two of the U5 snRNP proteins that later join the A2-1 complex via tri-snRNP are of particular importance: Brr2p and Snu114p. While another U5 protein (Prp8p) is thought to act at the catalytic core of the spliceosome (Sontheimer 2001), these two proteins are responsible for the unwinding of U4 and U6 to allow for U6 to address the 5' splice site. In this function, the DEAD box helicase Brr2p is thought to act as the molecular motor to unwind the base pairing (Xu, Nouraini, Field, Tang, and Friesen 1996)(Lauber, P. Fabrizio, Teigelkamp, Lane, Hartmann, and R. Lührmann 1996)(D. H. Kim and Rossi 1999). In so doing, Brr2p appears to be regulated by regulatory G protein Snu114p (Small, Leggett, Winans, and Jonathan P Staley 2006). From this standpoint, these two proteins are responsible for the events that allow the spliceosome to progress to the activated (A1) spliceosome.

Concurrent to the actions of Brr2p and Snu114p is the unwinding of U1 from the 5' splice site. This ATP dependent step is executed by the DEAD box helicase Prp28p (J P Staley and C Guthrie 1999). This protein has been identified as part of the U5 snRNP (S. W. Stevens, Barta, H. Y. Ge, R. E. Moore, M. K. Young, T. D. Lee, and J. Abelson 2001). In spite of the clear requirement for Prp28p in the A2-1 complex, this protein has paradoxically not been identified in purifications of the U4/U6·U5 tri-snRNP (A. Gottschalk, G. Neubauer, J. Banroques, Mann, R. Lührmann, and P. Fabrizio 1999)(S. W. Stevens et al. 2001). The purification of the penta-snRNP complex (see section 1.1.3.6) did, however, contain Prp28p leading to the speculation that previously purified tri-snRNP complex exists as either a nonfunctional intermediate or artifact of high salt biochemistry (S. W. Stevens et al. 2001).

1.1.4.5 The A1 Complex

After the exit of U1 and U4, the A1 complex is formed. This complex finds a bound U6 at the intron of the 5' splice site. U6 is concurrently paired to the U2 snRNA, which remains bound near the branchpoint. The U5 snRNA is basepaired with the exon at the 5' splice site (Fantes and Jean Beggs 2000).

Prp19p and associated proteins (Cef1p, Prp46p, Syf1p, SNT309p, Isy1p, Syf2p, CWC2p) are collectively known as the Nineteen Complex (NTC) (Chun-Hong Chen, Wan-Chin Yu, Twee Y. Tsao, Lian-Yung Wang, Hau-Ren Chen, Jui-Yen Lin, Wei-Yu Tsai, and Soo-Chen Cheng 2002). The NTC is required for the stable adherence of the U5 and U6 snRNPs within the spliceosome (Chan and Soo-Chen Cheng 2005). As part of this function, the NTC defines the interaction of the U5 and U6 snRNAs with the 5' splice site and therefore is a requirement for a properly formed A1 complex. Although it is believed that the NTC joins the spliceosome around the time of U4 exit, the order of these events was unknown. The purification and characterized of the A2-1 complex (Chapter 4 of this work) sheds light on these events.

1.1.4.6 Transitions to the A2-2 Complex

The transition from the activated A1 spliceosome to the A2-2 complex requires the activity of the DEAH-box protein Prp2p (King and J.D. Beggs 1990). This ATP dependent action moves the 2' hydroxyl of the branchpoint adenosine into proximity with the phosphodiester bond that composes the 5' splice site (S.H. Kim and R.J. Lin 1996).

1.1.4.7 The A2-3 Complex and Beyond

After the branchpoint moves into the vicinity of the 5' splice site, the first step of splicing is initiated – although the exact trigger for this action remains unknown (Fantes and Jean Beggs 2000). At this point, the substrate mRNA exists as the 5' exon and a

lariat intermediate. The formation of the A2-3 complex is not only defined by the formation of substrate intermediate, but also by another conformational change within the spliceosome.

In order to prepare for the second step of splicing, the substrate mRNA must undergo a rearrangement within the spliceosomal machinery. This rearrangement moves the 5' splice site into close proximity with its 3' counterpart. Both the 5' and 3' splice sites are now bound by the U5 snRNA (Umen and C. Guthrie 1995). The rearrangement within the spliceosome is the result of the ATP dependent DEAH-box helicase Prp16p (Zhou and Reed 1998).

In addition to the requirement for Prp16p, additional protein components are required for transition to the A2-3 complex. On the other hand, while they are essential, the functions of Prp17p, Prp18p, and Slu7p remain less well characterized (Fantes and Jean Beggs 2000).

After the formation of the A2-3 complex, the second transesterification reaction occurs rapidly (Fantes and Jean Beggs 2000). Later actions of Prp22p and Prp43p serve to release the spliced mRNA (J.D.O. Wagner, Jankowsky, Company, Pyle, and J.N. Abelson 1998) and lariat intron (Martin, S. Schneider, and Schwer 2002), respectively. The lariat intron is debranched by Dbr1p (Chapman and Boeke 1991) and later degraded.

1.1.4.8 The Mechanism of the mRNA Splicing Reaction

It is common speculation that at the core, the spliceosome is a ribozyme (Sontheimer 2001). Like most RNA enzymes, the spliceosome requires a divalent ion cofactor, likely Mg^{2+} (Villa, Pleiss, and C. Guthrie 2002)(Yean, Wuenschell, Termini, and R J Lin 2000). The resulting mRNA splicing reaction occurs in two steps. Both step one and step two are transesterification reactions (Alberts, Alexander Johnson, Lewis, Raff, Keith Roberts, and Walter 2002).

1.1.4.8.1 Chemistry of the First Step

Step one (Illustration 1.2) of the splicing reaction begins with the attack of the 2' hydroxyl group oxygen of the branchpoint adenosine on the phosphate at the 5' end of the intron (Berg, Tymoczko, and Stryer 2006)(Illustration 1.2B). The Mg^{2+} cofactor is thought to stabilize the leaving group intermediate in the reaction (Sontheimer 2001). This results in the cleavage of the phosphodiester bond between the 5' end of the intron and the 3' end of the upstream exon. Resulting from this process, the branchpoint adenosine and 5' terminal phosphate are linked by a 2',5' phosphodiester bond (Illustration 1.2C). This molecule is now known as the lariat intermediate (Domdey, Apostol, R.J. Lin, Newman, Brody, and J. Abelson 1984).

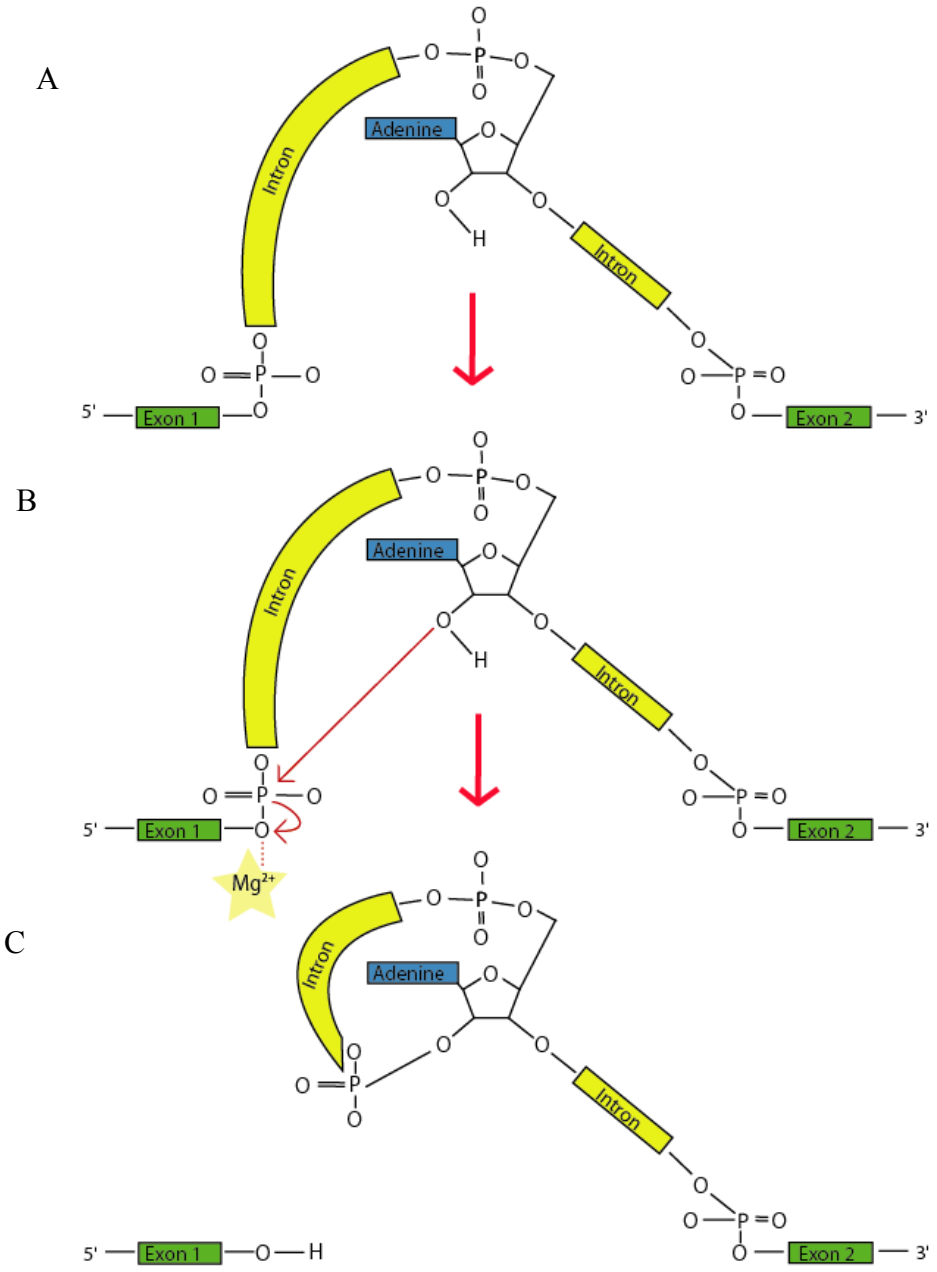


Illustration 1.2 – The first transesterification reaction of mRNA splicing

Figure adapted from (Darnell, Matsudaira, Zipursky, Lodish, Berk, and Baltimore 1999)

1.1.4.8.2 Chemistry of the Second Step

The second step of the splicing reaction starts with the intermediate A2-2 complex (Illustration 1.3A). With the conformational change to the A2-3 complex, the 5' and 3' splice sites are moved into close proximity. The hydroxyl group located on the 3' end of the 5' exon attacks the phosphodiester bond between intron and 3' exon (Illustration 1.3B). This results in a completely excised lariat intron and linked exons (Illustration 1.3C). The splicing reaction is now complete, and the component snRNPs are recycled.

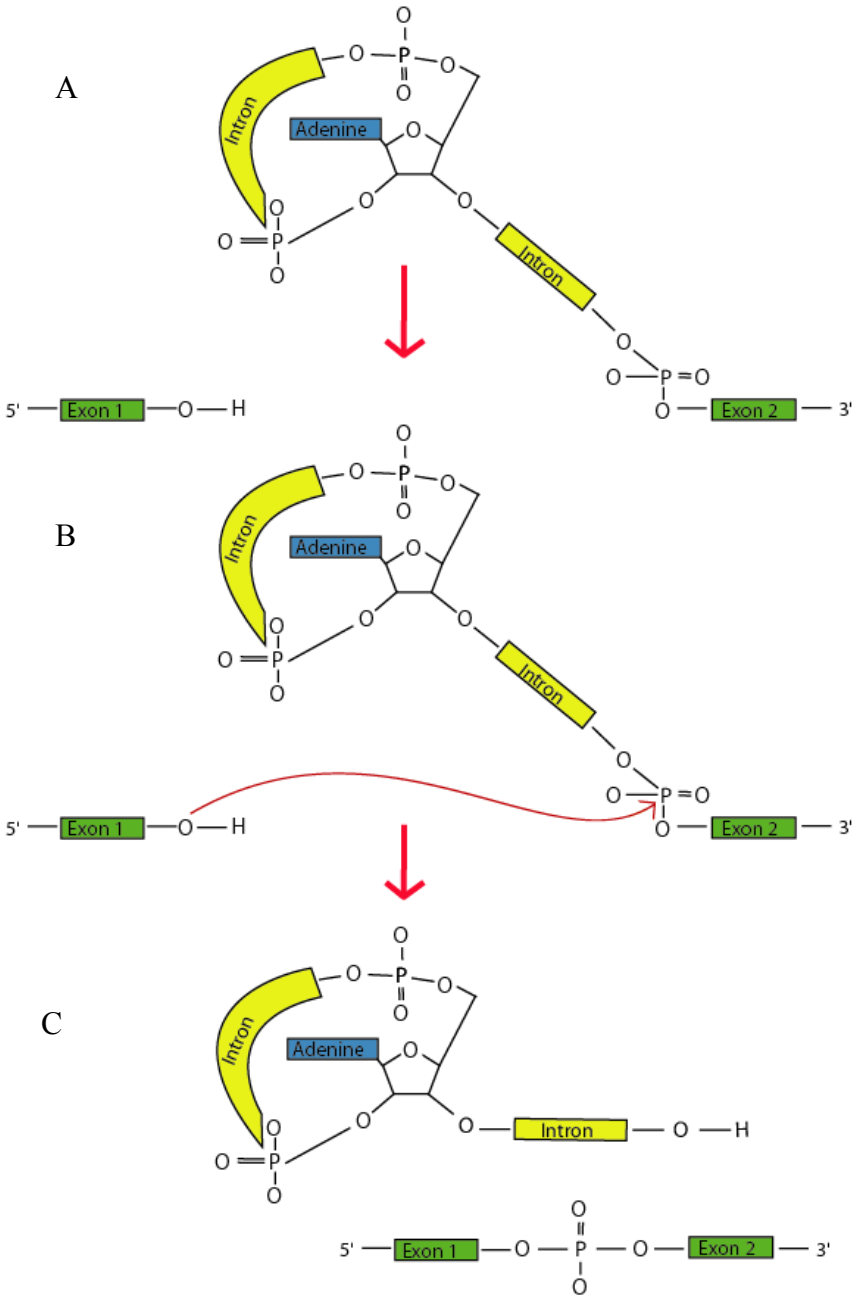


Illustration 1.3 – The second transesterification reaction of splicing

Figure adapted from (Darnell et al. 1999)

1.1.4.9 Assembly Models: Stepwise Versus Preformed

Although the model presented in Illustration 1.1 is based on a stepwise association of snRNPs to the pre-mRNA, there has recently been some debate as to the accuracy of this paradigm. Specifically, significant data exists that the spliceosome may assemble in absence of a substrate (Scott W. Stevens et al. 2002). Under this hypothesis, the progression of steps seen in Illustration 1.1 are not the result of component entry and exit, rather, each step simply represents a rearrangement of preexisting components.

The preformed model is the result of data indicating that spliceosomal component stability is highly salt sensitive. While characterizing the tri-snRNP (Scott W. Stevens and John Abelson 1999) under 250mM KCl purification conditions, Stevens et al noted that decreased salt concentrations of 150mM resulted in the purification of a 30S U2·U4/U6·U5 tetra-snRNP (unpublished data, noted in (Scott W. Stevens et al. 2002)). This work echoed previous findings in the mammalian system (M.M. Konarska and P.A. Sharp 1988) where the pre-mRNA free particle was termed the pseudospliceosome. Pursuing a further reduction in KCl concentrations to the physiological level of 50mM, Stevens and associates found that in addition to the tetra-snRNP, a 45S particle consisting of all five snRNAs was purified. Further, this particle was found to be missing the pre-mRNA binding proteins indicative of a substrate.

This data led to the alternative hypothesis that the spliceosome is a 45S RNP that is comprised of all five snRNAs. This complex is preformed in absence of a substrate pre-mRNA. In the event that a singular U1 snRNP should commit the pre-mRNA to splicing, a tetra-snRNP is recruited to form a functional A2-1 complex. Figure 1.4 represents this model.

In the period after the genesis of this model, there has been data both confirming and conflicting with the penta-snRNP model. Data from the mammalian system not only confirms the existence of a preformed spliceosome (Malca, Shomron, and Ast 2003)(Azubel, S.G. Wolf, J. Sperling, and R. Sperling 2004), but builds upon this model with the formation of a supraspliceosome (Azubel, Habib, R. Sperling, and J. Sperling 2006). This complex, which sediments at 200 S (Svedbergs) in a glycerol gradient, complex is thought to result from the engagement of four preformed spliceosomes onto a substrate mRNA. The supraspliceosome model was later confirmed by work in both HeLa cells as well as the DT40 chicken model (Yen-I G Chen, Roger E Moore, Helen Y Ge, Mary K Young, Terry D Lee, and Scott W Stevens 2007).

On the other hand, there have been a number of publications supporting the stepwise model of assembly. An analysis of cotranscriptional spliceosome formation in *Saccharomyces cerevisiae* revealed that a perturbation of the cap binding complex results in the spliceosomal loss of U5 on some transcripts while U1 andU2 snRNPs are retained (Görnemann, Kotovic, Hujer, and Karla M Neugebauer 2005). This data was later replicated using chromatin immunoprecipitation to address the spliceosomal assembly (Tardiff and Rosbash 2006). Lastly, analysis of HeLa extract indicates that U1 and U2 are able to form functional pre-spliceosomes independently of the U4/U6·U5 tri-snRNP or indeed the penta-snRNP complex (Behzadnia, Hartmuth, Cindy L. Will, and Reinhard Luhrmann 2006).

It is perhaps noteworthy that none of the findings have formally ruled out either pathway as inviable. Nor has the kinetic favorability of each respective model been addressed. In the end, it is a possibility that spliceosome assembly is achievable under both models, with one being the favored on pathway route under *in vivo* conditions.

For the sake of simplicity, I will use the traditional model of stepwise assembly throughout the remainder of this work to address the background of spliceosome assembly. It is therefore assumed that the preformed model is equally possible in these presentations.

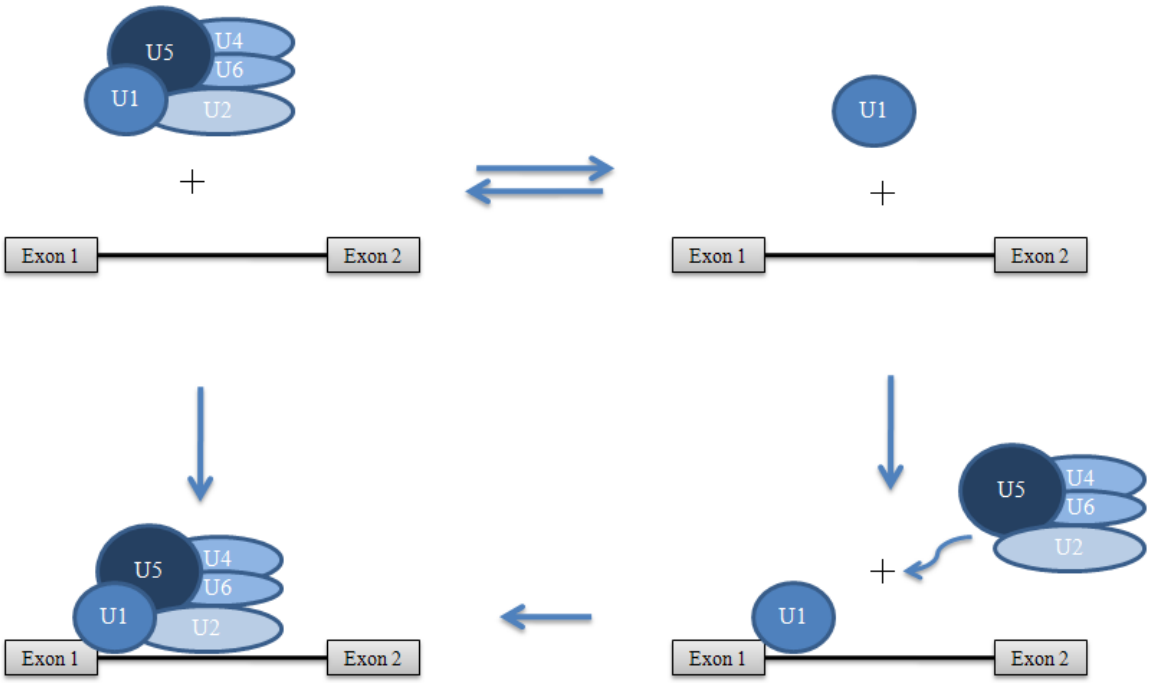


Illustration 1.4 – The penta-snRNP model of spliceosome assembly

Adapted from (Scott W. Stevens et al. 2002).

1.2 SACCHAROMYCES CEREVISIAE AS A MODEL ORGANISM

1.2.1 General

1.2.1.1 History

It can be said that yeast are the original biotechnology product as humans have been employing various species for more than 8,000 years in the production of beer (Thomas 2007). In spite of their widespread use, it wasn't until Anton van Leeuwenhoek first described yeast cells the late 17th century that their existence became known (Kornberg 1989). The link between the biochemical process of fermentation and yeast wasn't made until Pasteur's seminal work "Etudes sur la Biere" was published in 1876 (Baron 1996). Finally, in 1897, Eduard Buchner described the successful fermentation of a cell-free yeast extract. In many ways, this work represents the birth of modern biochemistry (Kornberg 1989). In the years following this discovery, the enzymes responsible for this activity were isolated. Although utility was found in other model organisms, yeast has retained its position as one of the preeminent models for study. Although there are more than 1500 known species of yeast, brewer's yeast or *Saccharomyces cerevisiae* remains the most widely characterized.

Over the past century, numerous biochemical and genetic techniques have been developed for use with yeast. No doubt a result of this toolkit, the complete genome of *S. cerevisiae* was published in 1996 (Goffeau et al. 1996). As the first eukaryote to be fully sequenced, *S. cerevisiae* once again proved its importance as the preeminent model organism.

1.2.1.2 Advantages

Yeast has a rapid doubling time, on order of two hours for a wildtype strain. It is relatively inexpensive to maintain and grow, and is easily archived in glycerol at -80°C. *Saccharomyces cerevisiae* readily undergoes homologous recombination, thus allowing for the insertion of exogenous sequences with ease (D Botstein and GR Fink 1988). The selection of these insertions is made uncomplicated by the existence of a chemically defined media and library of auxotrophs (Sherman 1991). A haploid phase allows recessive alleles to be altered without the complications found in diploid species. All of these features have resulted in an abundance of knowledge about this biological system and the techniques pertinent to it. The full sequencing of the genome revealed that the system operates on a modest 6000 genes, with none of the complications that result from alternative splicing.

1.2.1.3 Disadvantages

The disadvantages of *Saccharomyces cerevisiae* as a model organism all spring from its relative simplicity and evolutionary distance from humans. Being a single cell organism, it has inherent limitations in uncovering discoveries relevant to the human model. Nevertheless, the prime disadvantage of *Saccharomyces cerevisiae* as a model system for the study of RNA splicing, however, stems from the limited activity of this system. Perhaps due to the relative modest number of introns present in the organism, the number of active spliceosomes is somewhat limited in comparison to mammals. The result of this is that a fair amount of cell mass is needed for biochemical analysis, often requiring agitated fermentation. Nevertheless, in spite of the challenges of the system the advantages of *Saccharomyces cerevisiae* have made it the premier model for the study of mRNA splicing.

1.2.2 The Yeast Spliceosome as a Mammalian Model

Of the six thousand genes in *Saccharomyces cerevisiae*, only 283 contain introns. As there are only 296 introns present in the genome, the vast majority of genes that are spliced contain only a single intron of on average less than 500 nucleotides in size (Parenteau et al. 2008). These introns seem to be nonrandomly distributed, with 70% of ribosomal protein encoding genes undergoing splicing (Fantes and Jean Beggs 2000). Paradoxically, the vast majority of yeast introns appear to be nonessential for growth. A recent study found that the deletion of 96 separate introns (33% of all yeast introns) led to lethality in only five cases (Parenteau et al. 2008).

In contrast to yeast, the average human gene contains nine exons. Internal exons are 145 nucleotides long on average, with introns that are a mean of 3300bp long (Lander et al. 2001). Moreover, mammalian introns are known to contain biologically active elements like snoRNAs (Small nucleolar RNAs) within (Parenteau et al. 2008).

Perhaps owing to the increased complexity requirements, the mammalian system has been found to have two distinct spliceosomal systems (Woan-Yuh Tarn and Joan A. Steitz 1997). In addition to the traditional system composed of the U1, U2, U4, U5, and U6 snRNPs, mammals have been found to contain an alternative system referred to the ATAC spliceosome (Patel and Joan A Steitz 2003). This alternative system contains U11, U12, $U4^{ATAC}$, and $U6^{ATAC}$ while sharing U5 with the traditional U2-specific system (Luo, Moreau, Levin, and M J Moore 1999). (As this focus of this research is on the yeast spliceosomal system, all comparisons to the mammalian system will be with the traditional U2 spliceosome to the exclusion of the ATAC spliceosome.)

At the top level, the differences between mammalian and *Saccharomyces cerevisiae* splicing appear to be vast. Nevertheless, decades of work has revealed a great deal of conservation in the spliceosome itself. While and all five snRNA homologs were

identified early (C Guthrie and Patterson 1988), the subsequent years have revealed an array of proteins serving similar functions.

1.2.2.1 mRNA

Sequence conservation is much more strict in *Saccharomyces cerevisiae* than in mammals (Fantes and Jean Beggs 2000). With the exception of the leading nucleotides GU, the mammalian 5' splice site intron sequence is variable. The yeast sequence is more consistently a GUAUGU sequence (Sunbin Liu 2005). The branchpoint sequence is similarly comparable, with a highly conserved yeast sequence and a more variable mammalian. The 3' splice site sequence is rather short in both cases.

Overall, the greater variability of splice sequence in mammals is thought to enable alternative splicing (Fantes and Jean Beggs 2000). As *Saccharomyces cerevisiae* has no such requirements, a strict consensus splice site signal is probably favorable, both energetically and evolutionarily.

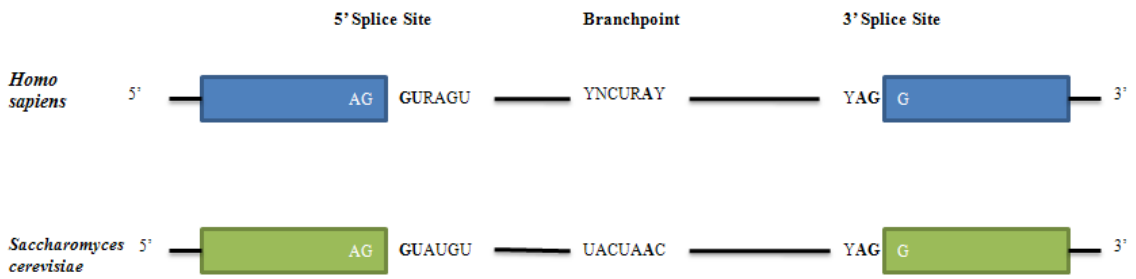


Illustration 1.5 – *S. cerevisiae* and metazoan consensus splice sequences

R represents any purine, while Y represents any pyrimidine. Sequences in bold are almost fully conserved. Adapted from (Sunbin Liu 2005)(Darnell et al. 1999).

1.2.2.2 snRNA

Excepting U6, snRNA length and primary sequence is widely divergent across species (Fantes and Jean Beggs 2000). On the other hand, the U6 sequence is highly conserved (Fantes and Jean Beggs 2000). The yeast U5 is alternatively processed at the 3' end to result in two final, active versions of the snRNA (Chanfreau, S A Elela, Ares, and C Guthrie 1997). The purpose of these two forms has not been elucidated to date. While the difference in size between yeast and humans is particularly large in the U1 and U2 snRNAs, the yeast U1 snRNA specifically contains additional domains that have no mammalian counterpart. It has been noted that the yeast U1 snRNP contains seven proteins that have no known mammalian counterpart. The presumption is that the existence of extra domains in the snRNA and the additional proteins are linked to U1 functionality in yeast that the mammalian counterpart lacks. Although the function of these players remains unproven, speculation exists that they are responsible for 5' splice site selection in an organism that does not require the flexibility of SR protein based enhancer selection (Fantes and Jean Beggs 2000).

In spite of the differences in primary sequence, the snRNA secondary structure is well conserved between yeast and mammals (Birnstiel 1988). The polymerase II products of both species contain Sm binding domains, and similarly contain Lsm binding regions in the RNA polymerase III produced U6.

	Saccharomyces cerevisiae		Homo Sapiens
snRNA	Gene	Size (nt)	Size (nt)
U1	SNR19	568	164
U2	SNR20	1175	187
U4	SNR14	160	145
U5S	SNR7	179	-
U5L	SNR7	214	116
U6	SNR6	112	106

Table 1.3 - Comparison of yeast and mammalian snRNAs

In every instance, the snRNAs of *Saccharomyces cerevisiae* are larger than their mammalian counterparts (table adapted from (Fantes and Jean Beggs 2000) and (Krainer 1997)).

1.2.2.3 Protein

The first ten snRNP proteins were purified from rat tissue in 1975. Since this time, scores of proteins have been identified that are either essential or advantageous for spliceosome function. Nevertheless, it was in probing these functional questions that the limitations of mammalian systems became apparent. As noted above in section 1.2.1.2, *Saccharomyces cerevisiae* as a model organism presents a number of advantages with regard to the practice of genetics. Ease of creating gene knockouts, epitope tagging, and rapid doubling time resulted in the ability to probe genetic interactions that weren't possible in mammalian counterparts.

Over time, spliceosomal proteins have been identified in both systems. Many of these have been found to be functional and structural homologs. Unfortunately, the protein naming conventions do not make these partners in homology readily apparent. To ease in discussion of these systems in later chapters, a table of known homologs is included in Appendix I.

1.3 AFFINITY PURIFICATION AND SEPARATION METHODS

1.3.1 The Tandem Affinity Purification

The *Staphylococcus aureus* protein-A has long been utilized in the laboratory to purify IgG antibodies and associated antigens (Kessler 1975). This protein binds antibodies rapidly with high affinity. While protein-A fusions were in use prior to the late 1990's (Popplewell, Gore, Scawen, and Atkinson 1991)(Stirling, Petrie, Pulford, Paterson, and Michael J. R. Stark 1992), it wasn't until 1999 that this tool was widely utilized. The advent of the Tandem Affinity Purification (TAP) tag allowed for the two step purification of protein complexes under native conditions (Rigaut, Shevchenko, Rutz, Wilm, Mann, and Séraphin 1999). Whether the tag sequence is attached to the C (Rigaut et al. 1999) or N (Puig, Caspary, Rigaut, Rutz, Bouveret, Bragado-Nilsson, Wilm, and Séraphin 2001) terminus of the bait protein's genome sequence, options exist to insure accessibility and avoidance of steric interactions. The tag consists of a protein-A domain and a calmodulin binding domain separated by a Tobacco Etch Virus (TEV) protease cleavage site. This configuration allows a primary purification using IgG sepharose beads. The protein and associated complex is then eluted under identical buffer conditions through cleavage of the TEV protease at the target sequence (Dougherty, Cary, and Parks 1989). The second affinity step can then utilize the calmodulin binding domain (Niggli, Penniston, and Carafoli 1979) Washing the protein complex while bound to the calmodulin sepharose removes additional contaminants as well as the TEV protease used for elution from the first step. Elution from the calmodulin sepharose beads requires removal of free Ca^{2+} from the buffer solution. This is typically achieved through use of ethylene glycol tetraacetic acid (EGTA) as a chelating agent. It is noteworthy that

EGTA not only binds calcium, but also free magnesium. This is an important consideration when purifying RNPs that require Mg^{2+} for proper RNA folding as the removal of this ion often dissolves the particle.

1.3.2 The Glycerol Gradient

As one of one of the most established biochemical separation techniques (Brakke 1953), the density gradient has been well established as a prime method for separation of particles by size and mass. The sedimentation coefficient in Svedberg units (S) has been assigned to most cellular components, thus allowing for easy comparison of particle size (Berg, Tymoczko, and Stryer 2006). More recently, the glycerol gradient has gained favor in the separation of spliceosomal complexes. Established gradient conditions allow for the rapid separation of mono-snRNPs from the U4/U6 di-snRNP, and tri-snRNP (Scott W. Stevens and John Abelson 1999), while alternate conditions must be considered for fully assembled spliceosomes (Scott W. Stevens et al. 2002).

1.4 Mass Spectrometry as a Method of Identifying Proteins

The beginnings of mass spectrometry as a field of research stretch back over a century (Thompson 1907). As an analytical technique, it found use beginning in 1919 determining atomic weights (Borman, Russell, and Siuzdak 2003). Due to difficulties in ionizing highly polar peptides, it wasn't until the 1980's that the field was sufficiently advanced for use in the biosciences (Mann, Hendrickson, and Pandey 2001). The development of matrix-assisted laser desorption/ionization (MALDI) and electrospray ionization (ESI) along with automated spectra analysis enabled the use of the mass spectrometry to identify short peptide sequences.

A combination of technologies is utilized in the gathering of mass spectrometric data. These technologies can broadly be grouped into two categories: ionization method

and the mass spectrometric instrumentation itself. While there are two dominant techniques for ionization of peptides, there are multiple types of instrumentation in popular use. Formally all three of these instrumentation techniques could be used independently with either ionization technique. In practice, however, MALDI is usually paired with time of flight (TOF) analysis. Likewise, ESI has been popularly paired with quadrupole and ion-trap spectrometers.

Matrix-assisted laser desorption/ionization was paired with time of flight analysis (MALDI-TOF) to create a system that ionizes sample via laser excitation then identifies the mass to charge ratio based on the time of flight to a detector (Karas and Hillenkamp 1988). Once created, the ions are driven to the detector through use of an acceleration field by way of an ion mirror that corrects for differences in initial energy.

Like all techniques, MALDI-TOF has both advantages and disadvantages when compared to competing technologies (Carr, Hemling, Bean, and Gerald D. Roberts 1991). Among its strengths are the relative simplicity of the system and the insensitivity to salts. The disadvantage of low resolution limits is more often overshadowed by the inability to be easily paired with liquid chromatography.

At one time, multiple charge states complicated the interpretation of electrospray ionization data (Carr, Hemling, Bean, and Gerald D. Roberts 1991). As computational horsepower has increased this limitation has subsided and ESI has become the dominant ionization method used in peptide identification. This technique pumps the sample through a fine needle so that micrometer sized droplets are ejected with help of an inert, nebulising gas into the charged spectrometer field (Yamashita and Finn 1984). Once in flight, the ion is commonly analyzed by spectrometry in the quadrupole and ion trap configuration (Fenn, Mann, Meng, S F Wong, and Whitehouse 1989). Quadrupole detection results from the generation of an electric field from four rods. This field acts as

a mass filter, allowing for the detection of ions of a given mass. The scanning of the field then allows the detection of a range of masses (Mann, Hendrickson, and Pandey 2001). Ion trapping allows the stream of ions to be captured and then later ejected by the alteration of a magnetic field (Mann, Hendrickson, and Pandey 2001).

Although not strictly confined to use with ESI, tandem mass spectrometry has been paired with ESI to great effect (Biemann and Scoble 1987). This technique pairs two spectrometers together to achieve greater resolution than would be possible with a single detector. Briefly, after detection in the first spectrometer, molecules are routed into a collision chamber where they are fractured by contact with an inert gas. The fragments of the peptide are then identified in a second detector followed by computational analysis that cross references data from the primary detection with the secondary.

For numerous reasons, ESI-MSMS has become the most popular technique set for peptide analysis. One of the advantages of ESI is the ability to pair it with one or more liquid chromatography steps. This allows for the separation of a complex peptide mixture prior to introduction into the analytical equipment. The most promising variant to utilize ESI is termed Automated Ultra-High-Pressure Multidimensional Protein Identification Technology (UHP-MudPIT) (Motoyama, Venable, Ruse, and John R Yates 2006). This system analyzes a trypsinized protein sample via a two dimensional chromatography separation prior to introduction to ESI. The first separation relies on a strong cation exchange column. The second separation utilizes a reverse phase gradient. Detection relies on tandem mass spectrometric analysis.

1.5 DISSERTATION OBJECTIVES

The goals of this project were originally quite different than what the end result became. The project began as an attempt to prepare the U5 and tri-snRNPs for cryo-electron microscopy. To achieve this goal, the complex to be purified would need to be

completely homogenous. These *Saccharomyces cerevisiae* complexes contain two different isoforms of Snu40p and Snu66p, as well as the long and short versions of the U5 snRNA. With this in mind, I set out to create a double *Δsnu40 / Δsnu66* knockout strain. When this strain was created, however, I noted that the double deletion resulted in a slight growth defect at lower temperatures. Serial dilutions incubated at a range of temperatures later confirmed this defect. Later molecular analysis revealed the tri-snRNP formation and function defects that became the basis of this dissertation. As is so often the case in science, serendipity and vigilance ruled the day.

This work is divided into four chapters. Chapter one provides a synopsis of the spliceosome, its component parts, and the splicing reaction these components drive. The chapter also compares *Saccharomyces cerevisiae* as a model organism against the benchmark of mammalian cell culture. Lastly, the chapter reviews the methods of protein complex purification, separation, and mass spectrometric analysis.

Chapter two details the materials and methods used throughout this work. The chapter begins with a listing of plasmids, strains, DNA oligonucleotides, and equipment used generally throughout. The next two sections detail the methods and materials used specifically for the work in chapters three and four.

Chapter three details the effects of a deletion of *SNU40* and *SNU66* from the yeast genome. The section shows that a double deletion of *SNU40* and *SNU66* leads to synthetic lethality at 16°C. Later work shows the effect of these deletions, as a singularly and in concert, on tri-snRNP formation. Lastly, this section reveals that the double deletion strain creates tri-snRNP that is deficient in U4/U6 unwinding activity at low temperatures. This inability to unwind U4 from U6 likely stalls the spliceosome at the precatalytic stage and leads to the synthetic lethality seen at restrictive temperatures.

Chapter four builds upon the previous chapter by purifying and characterizing the precatalytic spliceosome. The double knockout mutants used in chapter four revealed a buildup of a large complex at restrictive temperatures. This chapter uses affinity purification to isolate this complex. Consistent with the precatalytic spliceosome, northern blotting identifies the complex as containing all five snRNAs. Mass spectrometric analysis reveals the identity of numerous protein components present. Lastly, it is revealed that this precatalytic spliceosome is capable of catalytic activity.

Taken together, these results achieve at least three objectives. Firstly, the function is determined for the nonessential proteins Snu40p and Snu66p. Secondly, the functional proof and importance of Prp6p as a bridging factor between the U4/U6 di-snRNP and the U5 snRNP is emphasized. Lastly, a catalytically active precatalytic spliceosome has been purified and characterized in *Saccharomyces cerevisiae*. The dataset resulting from this characterization has implications for numerous aspects of the splicing cycle.

Chapter 2: Methods and Materials

2.1 Plasmids Generally Used

pBS1479 (Puig et al. 2001) and pBS1365 (Puig et al. 2001) were used in the creation of TAP tagged strains. pMPY-2xMYC (B L Schneider, Seufert, Steiner, Q H Yang, and Futcher 1995) was employed to make gene knockout constructs. Both are described in greater detail in subsequent sections.

2.2 Yeast Strains Generally Used

Strain	Mating Type	Genotype
BY4734	α	his3Δ200 leu2Δ0 met15Δ0 trp1Δ63 ura3Δ0
ySS1034	α	his3Δ200 leu2Δ0 met15Δ0 trp1Δ63 ura3Δ0 + Δsnu40
ySS1036	α	his3Δ200 leu2Δ0 met15Δ0 trp1Δ63 ura3Δ0 + ΔSNU66
ySS1038	α	his3Δ200 leu2Δ0 met15Δ0 trp1Δ63 ura3Δ0 + ΔSNU40 + ΔSNU66
ySS2013	α	his3Δ200 leu2Δ0 met15Δ0 trp1Δ63 ura3Δ0 + PRP6-TAP::URA3
ySS2014	α	his3Δ200 leu2Δ0 met15Δ0 trp1Δ63 ura3Δ0 + ΔSNU40 + PRP6-TAP::URA3
ySS2015	α	his3Δ200 leu2Δ0 met15Δ0 trp1Δ63 ura3Δ0 + ΔSNU66 + PRP6-TAP::URA3
ySS2016	α	his3Δ200 leu2Δ0 met15Δ0 trp1Δ63 ura3Δ0 + ΔSNU40 + ΔSNU66 + PRP6-TAP::URA3
ySS2017	α	his3Δ200 leu2Δ0 met15Δ0 trp1Δ63 ura3Δ0 + BRR2-TAP::URA3
ySS2018	α	his3Δ200 leu2Δ0 met15Δ0 trp1Δ63 ura3Δ0 + ΔSNU40 + BRR2-TAP::URA3
ySS2019	α	his3Δ200 leu2Δ0 met15Δ0 trp1Δ63 ura3Δ0 + ΔSNU66 + BRR2-TAP::URA3
ySS2020	α	his3Δ200 leu2Δ0 met15Δ0 trp1Δ63 ura3Δ0 + ΔSNU40 + ΔSNU66 + BRR2-TAP::URA3

Table 2.1 – Yeast strains used throughout

2.3 DNA OLIGONUCLEOTIDES GENERALLY USED

Oligo Name	Sequence
U1	5'-GAATGGAAACGTCAGCAAACAC-3'
U2	5'-CTTAAAAAGTCTCTTCCCGTC-3'
U4	5'-ACCATGAGGAGACGGTCTGG-3'
U5	5'-ATGTTCGTTATAAGTTCTATAGGC-3'
U6	5'-AGGGGAACTGCTGATC-3'
RPS4A Exon1	5'-GCAAAGATGGCTAGAGGACC-3'
RPS4a Exon2	5'-TCGGTTCTAACCTTACCGTCC-3'
RPS4a Intron	5'-GAGCACAATGAAAAAAATAAAATG-3'
RPS10 Exon1	5'-GCCAAAGGAAGACAGAAACAAGATCCAC-3'
RPS10 Exon2	5'-TTGAGTCTTGACGTAACCCTTAG-3'
RPS10 Intron	5'-TTCAATTTCAACGGAGAGATGTAAAC-3'
RPL16A Exon1	5'-GTCTGTTGAACCAGTTGTTGTC-3'
RPL16A Exon2	5'-GGGTCTTATTGAAAGCGGTAGC-3'
RPL16A Intron	5'-GAATGAGCATCCATTCATTCATC-3'
Snu40KOA	5'-AAGAACAAGGGAAAACTTTGATAAAAACACTCAAGTGC ACCATATTGTTTGTAGAAGGATAGGGAACAAAAGCTGG-3'
SNU40KOB	5'-TAAACTTAGTACTATATAGTCTCGCAGGCAATTCTCTT GAATAATTTCAACTTTTATATCCTATAGGGCGAATTGG-3'
SNU40PCR3	5'-GTGTACAAGGAAACCAGGACAGTAA-3'
SNU40PCR5	5'-CTTGTTTGTGAATGAGGGTCCCTTCC-3'
SNU66KOA	5'-ATATTAATAAAAAGGCAATCATCACATCAA CCCTTAATTAATTTAGGGAACAAAAGCTGG-3'
SNU66KOB	5'-CCAATTCGCCCTATAGCCTAATCAGCAAA TTCTACTAGCCCCAGCTTTTAATATTCAATA-3'
SNU66PCR3	5'-GTCATTTCACAGCGAAGTGTCATT-3'
SNU66PCR5	5'-TACATAGTTTTAGGTACTTTATATC-3'
PRP6TAPA	5'-TGCAATACTGCACACCTAGAGAGATTTTATTGCG CTTGATGAATGACAAATCCATGGAAAAGAGAAG-3'
PRP6TAPB	5'-CCCTATAGTGAGTCGTACTATCTTGCGTTGTTTT CTACTTAGCCTGCGCGTATATTTATATGTATGT-3'
BRR2TAPA	5'-TCTTGACGCAGATAAAGAGTTGTCCTTTGAAA TAAATGTGAAATCCATGGAAAAGAGAAG-3'
BRR2TAPB	5'-CGCAAGAGAATGTTATATATTGAAATCCA TTCGATTATCCAGGTACGACTCACTATAGGG-3'

Table 2.2 – DNA oligonucleotides used throughout

2.4 EQUIPMENT GENERALLY USED

Item	Supplier
UV Lightbox	Alpha Innotech
Alphaimager Gel Documentation System	Alpha Innotech
Nanopure Diamond Ultrapure Water Filtration System	Barnstead
Spectrophotomer (du530)	Beckman Coulter
Avanti J-20 XP Centrifuge	Beckman Coulter
Optima LE-80K Ultracentrifuge	Beckman Coulter
PowerPak HC Gel Power Source	Bio-Rad
PowerPak 300 Gel Power Source	Bio-Rad
PowerPak 3000 Gel Power Source	Bio-Rad
Quantity One Image Analysis Software	Bio-Rad
Trans-Blot Cell Transfer Apparatus	Bio-Rad
Big Shot Hybridization Oven	Boekel Scientific
Centrifuge 5417C	Eppendorf
Centrifuge 5417R (Refrigerated)	Eppendorf
Mini vortexer	Fisher Scientific
Isotemp 202s water bath	Fisher Scientific
Clinical centriuge	Fisher Scientific
Typhoon phosphoimager	GE Healthcare
Miniplus Pump	Gilson
HE 33 Mini Horizontal Gel Box	Hoefer
Novex mini-cell vertical gel apparatus	Invitrogen
Freezer -30	Kenmore
Geiger Counter	Ludlum Instruments
Scale (delta range PG4002-S)	Mettler Toledo
AE163 Analytical Balance	Mettler Toledo
PTC-100 Thermocycler	MJ Research
ND-1000 Spectrophotometer	Nanodrop
Innova 4330 Refrigerated Incubator Shaker	New Brunswick Scientific
Innova 4430 Refrigerated Incubator Shaker	New Brunswick Scientific
MM300 Mixer Mill	Retsch
50ml Stainless Steel Canisters for Mill	Retsch
Freezer Ultima II -80	Revco
UV Stratalinker	Stratagene
Dynastir Stirrer	VWR
DS-500E Shaker	VWR
General Purpose Plate Incubator	VWR
PROTEAN II xi Gel System	Bio-Rad
iCycler Thermocycler	Bio-Rad
Gel Dryer SGD2000	Savant

2.5 Methods and Materials Specifically Used in Chapter 3

2.5.1 Yeast growth and plasmids

Yeast were cultured on YPD or dropout media as required (Sherman 1991). Plasmids used were pMPY-2xMYC (B L Schneider, Seufert, Steiner, Q H Yang, and Futcher 1995) for URA3 popout knockouts, pBS1479 for TAP-tagging (Puig et al. 2001), pTAP-URA3 was created by subcloning the NotI/ApaI fragment of pBS1365 (Puig et al. 2001) into NotI/ApaI digested pBS1479 to allow TAP tagging with a URA3 marker.

2.5.2 Generation of Knockout Strains

BY4734 (MATα, his3Δ200, leu2Δ0, met15Δ0, trp1Δ63, ura3Δ0) was used as the parent of all strains. All knockouts were created through use of a PCR-generated replacement fragment containing 50 bp. arms of homology to the desired knockout region and a URA3 marker bounded by direct repeats (B L Schneider, Seufert, Steiner, Q H Yang, and Futcher 1995). Ura+ colonies were screened by PCR to determine the correct targeting following transformation using the lithium acetate procedure (Gietz and Woods 2002). Correctly targeted strains were subjected to 5-FOA-containing media to remove the URA3 (Boeke, Trueheart, Natsoulis, and G.R. Fink 1987). Double deletions were generated by removal of the second gene following the 5-FOA treatment of the first.

2.5.3 Generation of Affinity Tagged Strains

BY4734, ySS1034 *(*Δ*snu40*), ySS1036 *(*Δ*snu66*), and ySS1038 *(*Δ*snu40/*Δ*snu66*) were used as parent strains for the creation of TAP tagged Prp6p strains using pTAP-URA3: ySS2013 (BY4734 *PRP6*-TAP::URA3), ySS2014 (ySS1034 *PRP6*-TAP::URA3), ySS2015 (ySS1036 *PRP6*-TAP::URA3), ySS2016 (ySS1038 *PRP6*-TAP::URA3). Similarly, these parent strains were used in the creation of Brr2p-TAP strains ySS2017 (BY4734 *BRR2*-TAP::URA3), ySS2018 (ySS1034 *BRR2*-TAP::URA3), ySS2019

(ySS1036 *BRR2*-TAP::URA3), and ySS2020 (ySS1038 *BRR2*-TAP::URA3). Properly TAP-tagged strains were confirmed using western blotting with peroxidase-antiperoxidase complex (Rockland).

2.5.4 Extract Preparation

All cultures were grown in a volume of 1.5 L of YPD at 31°C unless otherwise noted. Strains grown at cold temperatures were grown to OD600=0.5 before chilling in ice water and shifting to 19°C for additional growth. Cells were then harvested at an OD (A600) of ~ 1.5 prior to processing. Briefly, cells were washed once in ice-cold ultrapure water, then washed a second time in 35 ml of AGK buffer (10 mM HEPES pH 7.9, 1.5 mM MgCl2, 200 mM KCl, 0.5 mM DTT, 10% glycerol) (Ansari and Schwer 1995). After the second wash, cells from 1.5 L were resuspended in 7 ml of AGK and frozen dropwise in liquid nitrogen prior to storage at -80°C. Cells were disrupted in a Retsch MM300 mill in liquid nitrogen using five cycles of three minutes each at 10 Hz. After thawing, the extract was clarified by centrifugation at 30,000 x g for 30 min at 4°C. The extract was centrifuged a second time at 100,000 x g for 1 hour at 4°C. The extract was then dialyzed in twice in 2 L of Buffer D (20% glycerol, 20 mM HEPES pH 7.9, 0.2 mM EDTA, 50 mM KCI, 0.5 mM DTT) for 1.5 hours at 4°C. The extract was centrifuged for 10 min at 20,000 x g and frozen at -80°C.

2.5.5 Serial Dilutions

Cultures were grown to between OD (A600) ~1.5 and diluted to OD (A600) 1.0. Subsequent ten-fold serial dilutions were made using sterile water. All dilutions were made on synthetic complete medium + 2% glucose.

2.5.6 Native Gel snRNP Analysis

Extracts (3 μl) were loaded to 1.5 mm thick native gels (4% polyacrylamide [80:1], 50 mM Tris base, 50 mM glycine, 2 mM MgCl2) (Raghunathan and C. Guthrie 1998a). Gels were electrophoresed at 160 V for 6 hours at 4°C with recirculating buffer. Where indicated, extracts were supplemented on ice with ATP or ATP+GTP, to the indicated concentrations and incubated at 19˚C or 31°C for 10 minutes prior to loading.

2.5.7 Denaturing RNA Gel Analysis

RNA was electrophoresed at 15 mA for 3 hours on a 20x16x1 mm denaturing polyacrylamide gel (7% acrylamide [19:1] 8M urea, 1 X TBE) (S.W. Stevens and J. Abelson 2002). This gel was then used in Northern analysis, gel extraction, or ethidium bromide visualization, depending on the application.

2.5.8 Northern Blotting

The gel was transferred overnight at 10 V and 4°C onto Brightstar membrane (Ambion) in an electroblotting apparatus (BioRad) in 25 mM sodium phosphate pH 7.0. After UV crosslinking for 30 sec, the membrane was placed into church buffer ((G.M. Church and W. Gilbert 1984) [7% SDS, 500 mM NaPO4 pH 7.0, 1 mM EDTA]). Pre-hybridization was performed for a minimum of 30 minutes at 42°C with rotation. Hybridization with γ-32P radiolabeled oligonucleotides was performed overnight at 42°C. The membrane was then washed briefly with 6x SSC. Two 15 minute washes were then performed in 2x SSC containing 0.5% SDS prior to a short 6x SSC wash. The membrane was then exposed overnight to a Bio-Rad phosphorimaging screen. The screen was developed in a GE Healthcare Typhoon phosphorimager.

2.5.9 Western Analysis

Protein samples were electrophoresed on 1 mm 17 well Invitrogen NuPAGE gels. Overnight transblotting was performed at 25V at 4°C in transfer buffer [20% methanol, 3.5 mM SDS, 386.3 mM glycine, 239.4 mM Tris base] to a nitrocellulose membrane (BioRad). The sample was then blocked for 1 hour in BLOTTO [5% dry milk in PBST (phosphate-buffered saline containing 0.2% Tween-20)]. After two 5-minute washes in PBST, the membrane was probed with peroxidase anti-peroxidase complex (rabbit; Rockland) in BLOTTO. The membrane was then washed 4 times with PBST for 15 minutes each. Enhance chemiluminescent detection was performed to detect the western blot signal (Perkin Elmer)

2.5.10 U4/U6 Unwinding Assay

20 µl of the indicated extract was mixed with 10 µl of IgG sepharose [10µg/µl IgG] in 200 µl NET50 buffer [50 mM Tris-Cl pH 7.4, 0.05% NP-40, 50 mM NaCl](Raghunathan and C. Guthrie 1998a). The mixture was rotated for 1.5 hours at 4°C and washed 3 times with 500 µl NET50. Identically prepared "E" samples in Figure 3.5 were stripped of RNPs with 400 µl of 1% SDS at this stage to demonstrate the total amount of U4, U5, and U6 bound to the beads. To prepare the supernatant "S" samples, washed beads were incubated for 5 minutes in 30 µl of unwinding buffer [2.5 mM MgCl2, 3% PEG-8000, 60 mM potassium phosphate pH 7.5, 8% glycerol, 8 mM HEPES pH 7.9, 0.08 mM EDTA, 20 mM KCl, 0.2 mM DTT]. Either ATP or the combination of ATP and GTP were added to a final concentration of 2 mM. The samples were incubated for 5-minutes at 11°C. The supernatant "S" was removed and the remaining affinity matrix was washed once with 350 µl NET50. This wash was combined with the 30 µl unwinding buffer as the total "S" sample. The beads were then stripped of remaining RNPs with 400 µl 1% SDS and constituted the beads "B" sample. All samples were

phenol chloroform extracted and the nucleic acids precipitated with 2.5 volumes of 100% ethanol at -30°C overnight.

2.5.11 Whole-Cell Extract Gradients

Strains ySS2013, ySS2014, ySS2015, and ySS2016 were grown at 31˚C, chilled in ice water and shifted to 19˚C prior to harvesting as described above. Extracts from were made as described above. After thawing on ice, 80 µl of cell extract was added to 120 µl of buffer GG [20 mM HEPES pH 7.0, 100 mM KCl, 0.2 mM EDTA, 5 µM Leupeptin, 3 µM Pepstatin, 0.4 mM PMSF, 0.5 mM DTT]. The mixture was layered onto a 10%-30% glycerol gradient and sedimented at 4˚C for 18 hours at 103,847 x g in an SW41 swinging bucket rotor. The gradients were harvested in 400 µl fractions, which were phenol-chloroform extracted. RNA and protein fractions were then precipitated overnight at -30° (S.W. Stevens and J. Abelson 2002).

2.5.12 Prp6p-TAP Affinity Purification

Extracts from strains ySS2013, ySS2014, and ySS2015, were prepared as described above except that extracts were dialyzed in Buffer D-Light (8% glycerol, 20 mM HEPES-KOH, pH 7.9, 0.2 mM EDTA, 50 mM KCI, 0.5 mM DTT.) Dialysis was performed twice for 1 hour each resulting in 5 ml of extract, which was then incubated with 250 µl of IgG agarose beads for 30 minutes at 4°C with rotation. The affinity-purified material was washed with 100 ml of Buffer D-Light. The beads were transferred to a microfuge tube and incubated with 400 µl Buffer D-Light containing 6 µg TEV protease for 1 hour at 17°C with rotation. The beads were harvested by centrifugation and the supernatant was layered onto a 10%-30% glycerol gradient. This was sedimented at 169,044 x g for 14 hours at 4°C in an SW41 rotor prior to fractionation, RNA extraction

and precipitation. RNA samples were electrophoresed through a denaturing urea-PAGE gel and northern blotted (see above).

2.6 Methods and Materials Specifically Used in Chapter 4

2.6.1 Extract Preparation for use in Affinity Purification

1.5L cultures grown at 31°C were used unless otherwise noted. Cultures were grown to OD_{600}=0.5 before temperature shifting the incubator to 19°C for additional growth. Once the cultures reached an OD (A_{600}) of ~ 1.5, they were harvested and processed as has been documented elsewhere. Briefly, cells were initially washed in 4°C ultrapure water. A second wash in 35 ml of AGK buffer (10 mM HEPES pH 7.9, 1.5 mM $MgCl_2$, 200 mM KCl, 0.5 mM DTT, 10% glycerol) was then executed. Lastly, the cell pellet was resuspended in 7 ml of AGK and frozen in liquid nitrogen. Long-term storage occurred at -80°C. Cell disruption was executed in a Retsch MM300 mill with 50ml canisters immersed in liquid nitrogen using five cycles of three minutes each at 10 Hz. The powdered, lysed cell material was then stored up to one week at 80°C. After thawing the material quickly in room temperature water, the cell debris was pelleted by centrifugation at 30,000 x *g* for 30 min at 4°C. A second centrifugation step of the resulting supernatant at 100,000 x *g* for 1 hour at 4°C cleared the extract. Extract was dialyzed twice in 2 L of Buffer D-Light (8% glycerol, 20 mM HEPES pH 7.9, 0.2 mM EDTA, 50 mM KCI, 0.5 mM DTT) for 1.5 hours each at 4°C.

2.6.2 IgG-Protein A Purification

Once dialysis of the extract was complete, it was incubated with 200µl of IgG sepharose for 30 minutes at 4°C. This material was then washed in a column with 100ml of Buffer D-Light. After the wash, the material was moved to a microfuge tube and placed in 400µl buffer D-Light. Tobacco Etch Virus (TEV) Protease (Dougherty, Cary,

and Parks 1989) was used to elute bound material from the column and was incubated for 1 hour at 19°C.

2.6.3 Glycerol Gradient Separation

After the TEV incubation, the sepharose was pelleted and the supernatant was removed. The ~400µl of material was placed on a 10%-30% glycerol gradient (0.5mM DTT, 1µM Leupeptin, 1µM Pepstatin, 0.4mM PMSF, 50mM KCl, 20mM HEPES pH8.0, 1.5mM $MgCl_2$, 0.02% NP-40.) The gradient was run at 51,892 x g for 16 hours / 4°C. The gradient was then fractionated by hand from the top in 400µl increments. For Figure 4.2, these fractions were phenol chloroform extracted, and the aqueous and phenol layers separated. The nucleic acids in the aqueous layer were then ethanol precipitated. The phenol layer was subjected to the addition of 2.5 volumes of acetone in order to precipitate any proteins present (Figure 4.3).

2.6.4 Calmodulin Purification

For Section 4.2.6 and 4.2.7, a second affinity step was utilized to minimize contaminants. Glycerol gradient fractions 15 through 19 were pooled into a single 2ml volume. This volume was then adjusted to a 2mM concentration of CaCl2. 6ml of Calmodulin Binding Buffer (0.5mM DTT, 10mM tris-Cl pH8.0, 50mM KCl, 1mM magnesium acetate, 1mM imidazole, 2mM $CaCl_2$, 0.1% NP-40, 8% glycerol) was added to the gradient material for a total of 8ml. Incubation with 200µl of calmodulin beads (Stratagene) followed for a period of 30 minutes at 4°C.

2.6.5 Post Glycerol Gradient Concentration

The particle used in section 4.2.5 was concentrated. To create this concentrated particle, 1.5 ml of glycerol gradient material (~20% glycerol) was diluted with 4ml of Gradient Dilution Solution (Conditions identical to those used in glycerol gradients, but

with 5% glycerol) to create a ~10% final glycerol concentration. This 5.5ml of material was then placed in an SW-55 and centrifuged for 16 hours at 122,871 x g / 4°C. After pelleting the material present, the supernatant was briefly removed from the tube while the pellet was resuspended in 100µl of supernatant. This material was then used in splicing reactions as indicated below.

2.6.6 RNA Sample Preparation

After the addition of a 10% volume of sodium acetate and three volumes of ethanol, the RNA samples were stored over night at -30°C. The next morning the tubes were spun for 10 minutes at 20,800 xg at 4°C. After 2x washes in 70% DEPC treated ethanol, the RNA pellets were placed to dry on the benchtop. Pellets were resuspended in 10µl DEPC treated H2O.

2.6.7 Protein Sample Preparation

Upon addition of 2.5 volumes of acetone to the organic phase of a phenol chloroform reaction, the samples were placed at -30°C over night. A 10 minute, 20,800 xg spin at 4°C in a microfuge pelleted any proteins. After two washes in 80% ethanol, the protein pellets were left to dry on the benchtop. Samples used in protein gels were resuspended in protein loading buffer while samples used for mass spectrometry were shipped in dry ice as dry pellets.

2.6.8 Mass Spectrometry

Mass spectrometric analysis was completed through collaboration with the laboratory of John Yates at the Scripps Research Institute, La Jolla California. Briefly, after phenol-chloroform separation of proteins the sample was acetone precipitated by addition of three volumes ice cold acetone. The sample was then placed at -20°C overnight. The next day, the proteins were spun down for 10 minutes at 20,800 x g and

washed twice in 80% ethanol. After the final wash was removed, the pellet was left to dry on the benchtop. After ~20 minutes the microfuge tube containing the sample was closed and shipped over night to the Yates lab on dry ice.

Upon arrival at the Yates lab, the sample was resuspended in a urea mix to denature the proteins (Washburn, Wolters, and J.R. Yates 2001). The sample was then reduced (TCEP) and alkylated. After a digest with trypsin and Lys-C, the sample was quenched with formic acid and loaded to the mass spectrometric apparatus. The sample was then analyzed by Automated Ultra-High-Pressure Multidimensional Protein Identification Technology (UHP-MudPIT) (Motoyama, Venable, Ruse, and John R Yates 2006)(Emily I. Chen, Hewel, Felding-Habermann, and John R. Yates 2006) using a ThermoFinnigan LCQ ion source. Spectra were matched to the Saccharomyces Genome Database (Eurie L Hong et al. 2008) by the SEQUEST algorithm (Ducret, Van Oostveen, Eng, J R Yates, and Aebersold 1998).

2.6.9 Protein Gels

Protein samples were run on Invitrogen NuPAGE 4-12% Bis-Tris Gels. Gels were then fixed in a 50% methanol / 12% acetic acid solution, washed in double distilled H_20, and coomassie stained.

2.6.10 Denaturing RNA Gels

RNA containing sample was placed on a 20x16x1 mm denaturing polyacrylamide gel (7% acrylamide:bisacrylamide [19:1] 8M urea, 1 X TBE) and electrophoresed for 2 hours at 20 mA.

2.6.11 Native RNA Gels

Native RNA samples were loaded onto a 20x16x0.75 mm gel containing 9% 19:1 acrylamide:bisacrylamide and 1xTBE. The gel was electrophoresed at 25mA for 2.5

hours at 4°C. The resulting gel was then transferred to a membrane for Northern blotting as described previously.

2.6.12 Psoralen Crosslinking of RNA

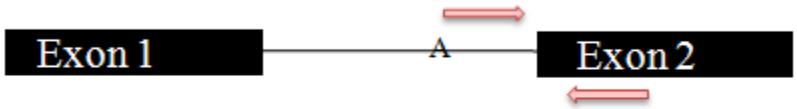
Glycerol gradient material (400µl) created as described in section 2.6.3 (Figure 4.2 pooled fractions 15-19) was Psoralen treated using methods described elsewhere (Yean and R.J. Lin 1991). Psoralen (Sigma-Aldrich) was added to a final concentration of 20µg/µl followed by UV (365nm, ~5cm distance from source) exposure on ice for 20 min. The samples were then phenol-chloroform extracted and precipitated as described in the RNA Sample Preparation section. Multiple samples of the RNA were run in multiple lanes. This gel was then split prior to membrane transfer, and three different membranes were transferred. These membranes were then exclusively probed for U2, U4, or U6.

2.6.13 RT-PCR

4µl of purified material from pooled glycerol gradient fractions 15-19 (Figure 4.2) was added to 6µl of H2O. This mixture was immediately treated with a 20% volume proteinase-K buffer (50mM EDTA pH 8.0, 1% SDS, 1mg/ml proteinase K) for 20 minutes at 40°C. Next, 200µl of RNA extraction buffer (50mM sodium acetate pH 5.3, 1mM EDTA, 0.1% SDS, 100 mg/ml glycogen) was added followed by 200µl of a 25:24:1 mixture of Phenol (pH 5.3) : Chloroform : isoamyl alcohol. Following two phenol chloroform extractions, the sample was chloroform extracted with 200µl chloroform. The aqueous layer was then treated with a 10% volume of 3M sodium acetate and ethanol precipitated with three volumes of ethanol. Samples were placed at -30°C over night. After centrifugation for 10 minutes at 20,800 x g the pellets were washed 2x with DEPC treated 70% ethanol. Finally, the pellet was left to air dry on the benchtop before resuspension in 10µl H2O. 1µl of this suspension was then used in an

RT-PCR reaction using the Quagen one step RT-PCR kit following the manufacturer's instructions. Primer sets against *RPS10* (Illustration 2.1), *RPS4A*, and *RPL16A* (not shown) were used in the RT-PCR reaction. Table 2.3 contains the program parameters used in the RT-PCR.

	RPS4A (bp)	RPS10A (bp)	RPL16A (bp)
Spliced	0	0	0
Full length	303	183	213



	RPS4A (bp)	RPS10A (bp)	RPL16A (bp)
Spliced	236	159	200
Full length	492	604	490

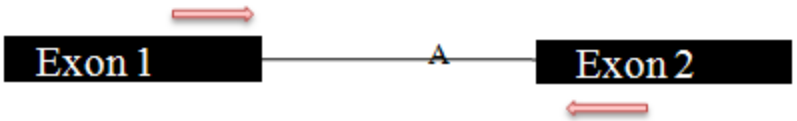


Illustration 2.1 – RT-PCR primer design

Step	Temperature (°C)	Time
1	55°	30 minutes
2	95°	15 minutes
3	94°	45 seconds
4	55°	45 seconds
5	72°	1min 45 seconds
6	Go to step 3, repeat 31 times	
7	72°	10 minutes
8	4°	Hold

Table 2.3 – RT-PCR parameters

2.6.14 Extract Preparation for use in Splicing Reactions

Whole-cell extract for use in splicing reactions was prepared in a manner similar to the extract used prior affinity purification. Briefly, the protocol was identical to the extract preparation described above, with the sole exception being a final dialysis step in Buffer D (20% glycerol, 20 mM HEPES pH 7.9, 0.2 mM EDTA, 50 mM KCl, 0.5 mM DTT) as opposed to the earlier Buffer D-Light. After dialysis of the extract, any remaining particulate was pelleted by microcentrifugation at 20,817 xg for 10 min at 4°C, The extract was then aliquoted, frozen in liquid nitrogen, and stored at -80°C.

2.6.15 Transcription of Radiolabeled mRNA

A plasmid was created by combination of the pUC19 vector and the actin gene sequence (*ACT1*). Production of the *ACT1* pre-mRNA was placed under the control of the

SP6 promoter. In order to prevent run-on transcription, the plasmid was linearized at the 3' end of the coding sequence through use of the EcoRI enzyme (New England Biolabs). Following digestion, the DNA was phenol chloroform extracted and ethanol precipitated. The transcription reaction itself was set up in the presence of α-P32 labeled GTP. After a two hour incubation at 40°C, the reaction was DNase I (1µl) treated for 15 minutes at 37°C. The reaction was then run on a denaturing RNA gel and the full-length RNA transcript was extracted.

2.6.16 Micrococcal Nuclease Treatment of Extract

BJ2168 whole cell extract prepared for use in splicing reactions was treated with micrococcal nuclease as described by Yean and Lin (Yean and R.J. Lin 1991). 1 µl of 20mM $CaCl_2$ was added to 15 µl of extract on ice. Finally, 4µl of micrococcal nuclease (Sigma 22 U/µl) was added and the reaction was incubated for 30 minutes at 30°C. The nuclease was quenched with the addition of 2 µl of 50mM EGTA and placed on ice until use.

2.6.17 Splicing Reactions

Splicing reactions were carried out as described previously (R.J. Lin, Newman, S.C. Cheng, and J. Abelson 1985) using either whole cell extract or extract that was micrococcal nuclease treated prior to use. Reactions were assembled on ice under the following buffer conditions: 60mM potassium phosphate pH 7.0, 2mM ATP, pH 7.0, 3.2mM MgCl2, 3% PEG-8000, 1mM spermidine. In addition to these additives, the 10µl reactions contained 1µl of radiolabeled pre-mRNA (~1,000 cpm/µl) and 4µl of either regular or micrococcal nuclease treated BJ2168 whole cell extract. Reactions were incubated at 22°C for 30 minutes. The reaction was then treated with a 2µl Proteinase-K buffer (50mM EDTA pH 8.0, 1% SDS, 1mg/ml Proteinase K) for 20 minutes at 40°C.

Next, 200 µl of RNA extraction buffer (50mM sodium acetate pH 5.3, 1mM EDTA, 0.1% SDS, 100 mg/ml glycogen) was added followed by 200 µl of a 25:24:1 mixture of phenol (pH 5.3) : chloroform : isoamyl alcohol. After a single repetition of the extraction, a 10% volume of 3M sodium acetate was added to the aqueous layer. The addition of three volumes of ethanol precipitated any nucleic acids present. Samples were placed at -30°C over night. After centrifugation for 10 minutes at 20,800 xg, the pellets were washed with DEPC treated 70% ethanol. The pellet was resuspended in denaturing loading buffer and run on a 20x16x1mm denaturing gel (7% 29:1 acrylamide: bisacrylamide, 7M Urea, 1X TBE) for 2 hours at 15mA. The gel was then dried and imaged in a phosphorimager (GE Healthcare Typhoon).

CHAPTER 3: SNU40P, AND SNU66P INTERACTION VIA PRP6P PROMOTES TRI-SNRNP

3.1 INTRODUCTION

The pre-mRNA splicing machinery requires rearrangement of multiple RNA and protein components prior to, during and after addressing the pre-mRNA substrate (J.P. Staley and C. Guthrie 1998)(Madhani and C. Guthrie 1994)(M.J. Moore, Query, and P.A. Sharp 1993). After the initial recognition of the 5' splice site by the U1 snRNP (Ruby and J. Abelson 1988)(B. Seraphin, Kretzner, and Rosbash 1988)(Siliciano and C. Guthrie 1988), the pre-mRNA branch point sequence interacts via RNA-RNA and RNA-protein contacts with the U2 snRNP in the first energy dependent step of the splicing reaction (J. Wu and Manley 1989)(Parker, Siliciano, and C. Guthrie 1987). This process is followed by the functional engagement of a pre-assembled U4/U6•U5 snRNP (M.M. Konarska and P.A. Sharp 1987)(S.C. Cheng and J. Abelson 1987). A number of structural rearrangements leading to formation of the activated spliceosome then occur. In the process, an intermolecular helix formed between U4 and U6 snRNAs is unwound by Brr2p, allowing the formation of a U2/U6 base-pairing interaction required for splicing (Raghunathan and C. Guthrie 1998b)(Hausner, Giglio, and Weiner 1990)(D.A. Wassarman and J.A. Steitz 1993)(Madhani and C. Guthrie 1992). After the catalytic steps of pre-mRNA splicing are completed, spliceosome disassembly liberates the remaining U2, U5 and U6 snRNPs from the lariat intron RNA (Martin, S. Schneider, and Schwer 2002)(Small, Leggett, Winans, and Jonathan P Staley 2006). U6 snRNP then base-pairs with U4 to form the U4/U6 di-snRNP (Raghunathan and C. Guthrie 1998a) which associates with the U5 mono snRNP to reform the U4/U6•U5 snRNP.

Many protein-protein interactions within the U4/U6•U5 snRNP have been previously described. Among the most well-studied interactions are between the U5 snRNP components Brr2p/[U5-200K], Prp8p/[U5-220K], and Snu114p/[U5-116K] (T. Achsel, Ahrens, Brahms, Teigelkamp, and R. Lührmann 1998). Other players in U4/U6•U5 snRNP formation include Prp6p/[U5-102K] a tetratricopeptide repeat (TPR) protein which is known to interact with the U5 proteins Snu40p/[U5-52K] (Laggerbauer et al. 2005), Dib1p/[U5-15K] (Uetz et al. 2000), Brr2p,Snu114p and Prp8p (S.B. Liu, Rauhut, H.P. Vornlocher, and R. Lührmann 2006) and the U4/U6•U5 snRNP-specific Snu66p/[U4/U6•U5-110K] (S.B. Liu, Rauhut, H.P. Vornlocher, and R. Lührmann 2006)(van Nues and Jean D. Beggs 2001). Additional interactions have been characterized between Prp6p and the U4/U6 components Prp3p/[U4/U6-90K] and Prp31p/[U4/U6-61K] (S.B. Liu, Rauhut, H.P. Vornlocher, and R. Lührmann 2006). Analysis of a prp6 mutant revealed that U4/U6•U5 snRNP formation is disturbed (Galisson and Legrain 1993). While yeast Prp6p is accepted to be stable components of the U4/U6•U5 snRNP, Prp6p has been previously characterized as a U4/U6 component (Abovich, Legrain, and Rosbash 1990) whereas the U5-102K protein is a stable component of the human U5 snRNP. These data have led to a model in which Prp6p acts as a bridging factor between the U5 snRNP and U4/U6 snRNP during the formation of the U4/U6•U5 snRNP, but the details of assembly have not been well characterized.

Among the partners known to interact with Prp6p are Snu40p and Snu66p; Snu40p is exclusively associated with the U5 snRNP in *Saccharomyces cerevisiae* (S. W. Stevens et al. 2001) while Snu66p is exclusively associated with the U4/U6•U5 snRNP (Scott W. Stevens and John Abelson 1999)(A. Gottschalk et al. 1999). The mammalian orthologs of Snu40p and Snu66p, U5-52K and U4/U6•U5-110K, are likewise found exclusively in the U5 and U4/U6•U5 snRNP respectively (Laggerbauer et al. 2005).

The GYF domain of U5-52K is thought to be responsible for protein-protein interactions (Freund, Dötsch, Nishizawa, Reinherz, and G Wagner 1999) and it has been crystallized in complex with one of its binding partners, U5-15K (Nielsen, Sunbin Liu, Reinhard Lührmann, and Ficner 2007). Deletion of *SNU40* results in no observable growth phenotype (S. W. Stevens et al. 2001) yet the yeast *SNU66* deletion mutant displays a slight cold sensitive growth phenotype and a splicing defect at reduced temperatures (S. W. Stevens et al. 2001).

We have determined that removal of SNU40 from *Δsnu66* strains resulted in synthetic lethality at reduced temperatures. In *Δsnu66* strains, U4/U6•U5 snRNP formation is inhibited, and at reduced temperatures, the U4/U6•U5 snRNP that has formed is defective for U4/U6 unwinding by Brr2p. In the *Δsnu40/Δsnu66* strain, the U4/U6•U5 snRNP biogenesis defect at reduced temperatures is bypassed, however the U4/U6•U5 snRNP remains defective in U4/U6 unwinding. Intriguingly, a transient Prp6p intermediate with U5 snRNP is reduced in *Δsnu40*, and increased in *Δsnu66*. We present a model in which U4/U6•U5 snRNP assembly proceeds through a series of events including serial interaction of Prp6p with Snu40p and Snu66p. Under unfavorable growth conditions, these non-essential proteins are required for biogenesis events to promote the correct assembly of the U4/U6•U5 snRNP.

3.2 RESULTS

3.2.1 Δsnu40 / Δsnu66 Strains are Synthetically Lethal at 16°C

Previous work has shown that *Δsnu66* cells are cold sensitive. We constructed isogenic *Saccharomyces cerevisiae* strains lacking *SNU40*, *SNU66*, or both. Serial dilutions of the strains show no growth phenotype for any deletion at 31°C, and the *Δsnu40* strain was slightly temperature sensitive at 37°C (Figure 3.1). Subsequent removal of SNU40 from *Δsnu66* resulted in synthetic lethality at 16°C (Figure 3.1) demonstrating a genetic interaction between these genes.

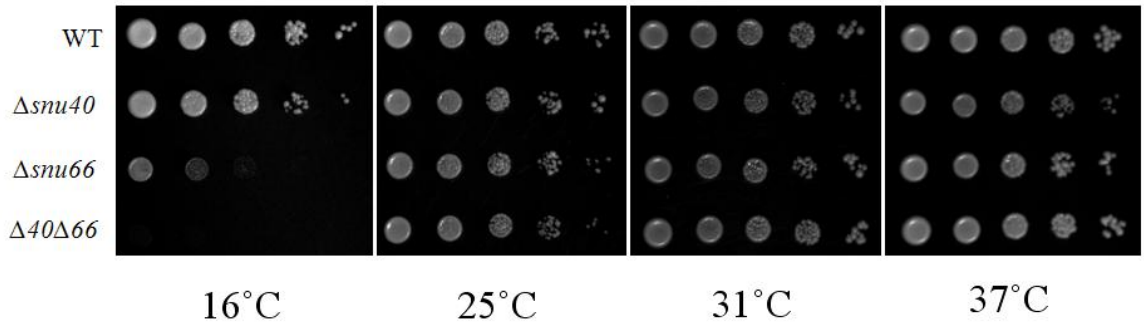


Figure 3.1 – *Δsnu4+Δsnu66* strains are synthetically lethal at 16°C

The cold sensitive Δ*snu66* strain is lethal at 16°C upon subsequent deletion of SNU40. Isogenic *SNU40/SNU66*, Δ*snu40*, Δ*snu66*, Δ*snu40/Δsnu66* strains were grown in YPD at 31°C, serially diluted and plated on SC media. Identically prepared plates were incubated at the indicated temperatures.

3.2.2 U4/U6 and U4/U6•U5 levels in Mutant Strains

Whole-cell splicing extract from *Δsnu40*, *Δsnu66*, and *Δsnu40/Δsnu66* and the isogenic wild type (WT) strain were generated and electrophoresed through a native gel designed to resolve the U4/U6•U5 snRNP from U4/U6 and U6 snRNPs (Raghunathan and C. Guthrie 1998a). Gels were transferred to membranes and northern blotted to detect the U6 snRNA. At 31°C, U4/U6•U5 snRNP levels were reduced in the Δsnu66 strains (Figure 3.2 lane 3) which was exacerbated in strains grown at 19°C (Figure 3.2 lane 7). The reduction in U4/U6•U5 snRNP levels at 19°C in *Δsnu66* was accompanied by an increase in U4/U6 indicating the defect was not in overall reduction in snRNP levels, but the inability to assemble U4/U6•U5 snRNP from U5 and U4/U6 snRNPs. The *Δsnu40* extracts do not show defects in snRNP assembly at either temperature (Figure 3.2 lanes 2, 6). Remarkably, in *Δsnu40/Δsnu66* extracts the U4/U6•U5 snRNP assembly defect seen in Δsnu66 extracts is bypassed by removal of *SNU40* (Figure 3.2 lanes 3, 4; 7, 8). We also note that the migration of the U4/U6 snRNP is slower in all strains lacking Snu40p (Figure 3.2 lanes 2, 4, 6, 8).

3.2.3 Reduced ATP-dependent U4/U6•U5 snRNP Dissociation

Whole-cell splicing extracts were supplemented with ATP to stimulate Brr2p-dependent unwinding of the U4/U6 snRNAs leading to U4/U6•U5 snRNP dissociation (Raghunathan and C. Guthrie 1998b)(S. W. Stevens et al. 2001). In the WT strain at 31°C, this resulted in reduced U4/U6•U5 levels (Figure 3.2 and Figure 3.3 lanes 1, 5, 9, 13) and increased levels of a U6 RNP lacking Prp24p, termed U6* (Raghunathan and C. Guthrie 1998b). The snRNP distribution in *Δsnu40* extracts, at both 31°C and 19°C in the

presence of ATP was indistinguishable from WT (Figure 3.3 lanes 10, 14). At 31°C in the *Δsnu66* strain, there is some dissociation of the U4/U6•U5 snRNP upon ATP treatment (Figure 3.2 lane 3, Figure 3.3 lane 11) but no dissociation at 19°C (Figure 3.2 lane 7, Figure 3.3 lane 15). In the *Δsnu40/Δsnu66* strain, a slight ATP-dependent disruption of the U4/U6•U5 snRNP was noted (Figure 3.2 lane 4, Figure 3.3 lane 12), however there is no disruption of U4/U6•U5 snRNP by ATP at 19°C (Figure 3.2 lane 8, Figure 3.3 Lane 16). We also note in the *Δsnu40/Δsnu66* extract, an additional U6 snRNP species was detected at permissive temperature (Figure 3.3 lane 12).

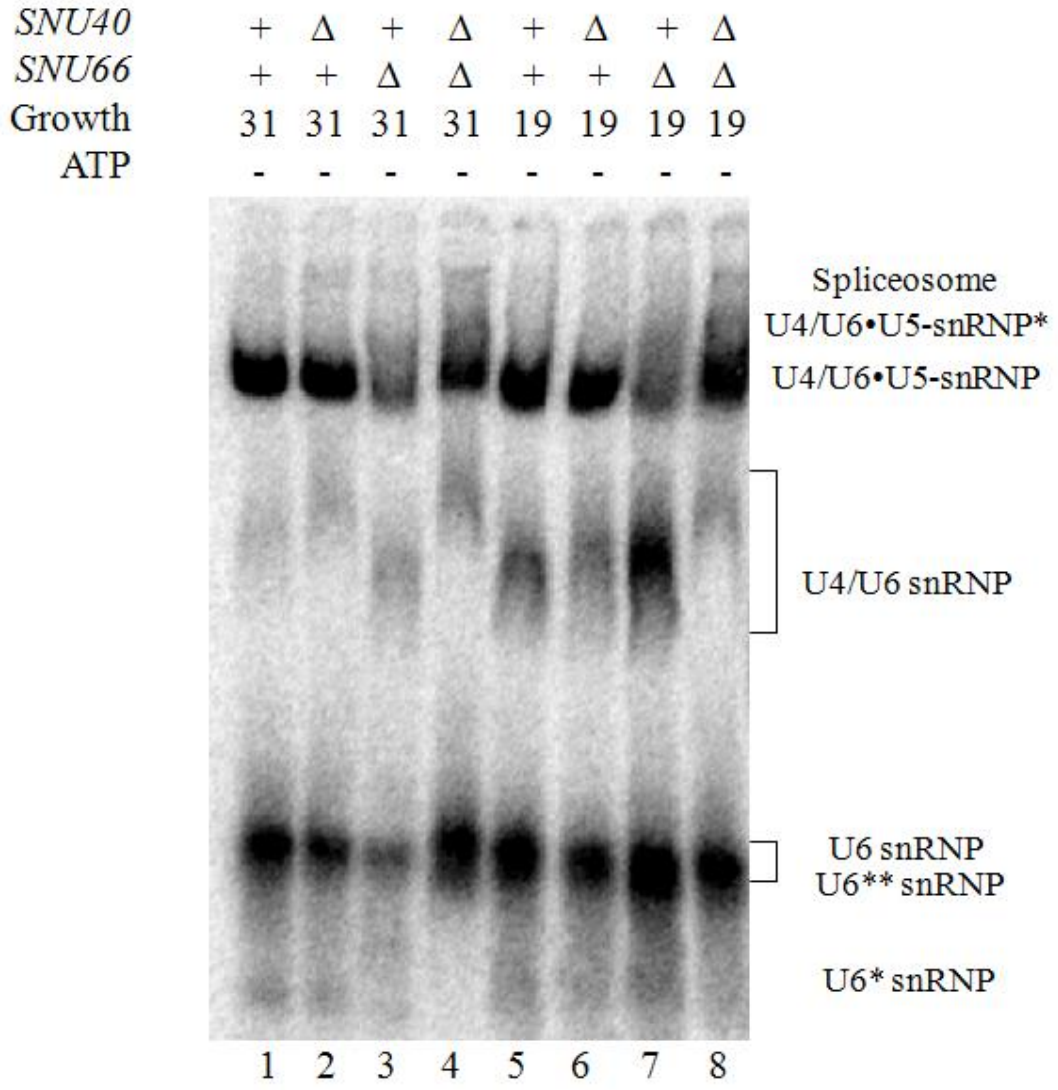


Figure 3.2 - U4/U6•U5 snRNP formation and dissociation in mutant strains

U4/U6•U5 snRNP formation is defective upon removal of *SNU66* and restored upon subsequent removal of *SNU40*. Cells were grown at 31°C and maintained at 31°C or shifted to 19°C until OD (A600) ~ 1.5. Splicing extracts were prepared and incubated at the indicated temperatures prior to electrophoretic separation on native polyacrylamide gels. U4/U6, U6, U6* and spliceosome complexes are indicated.

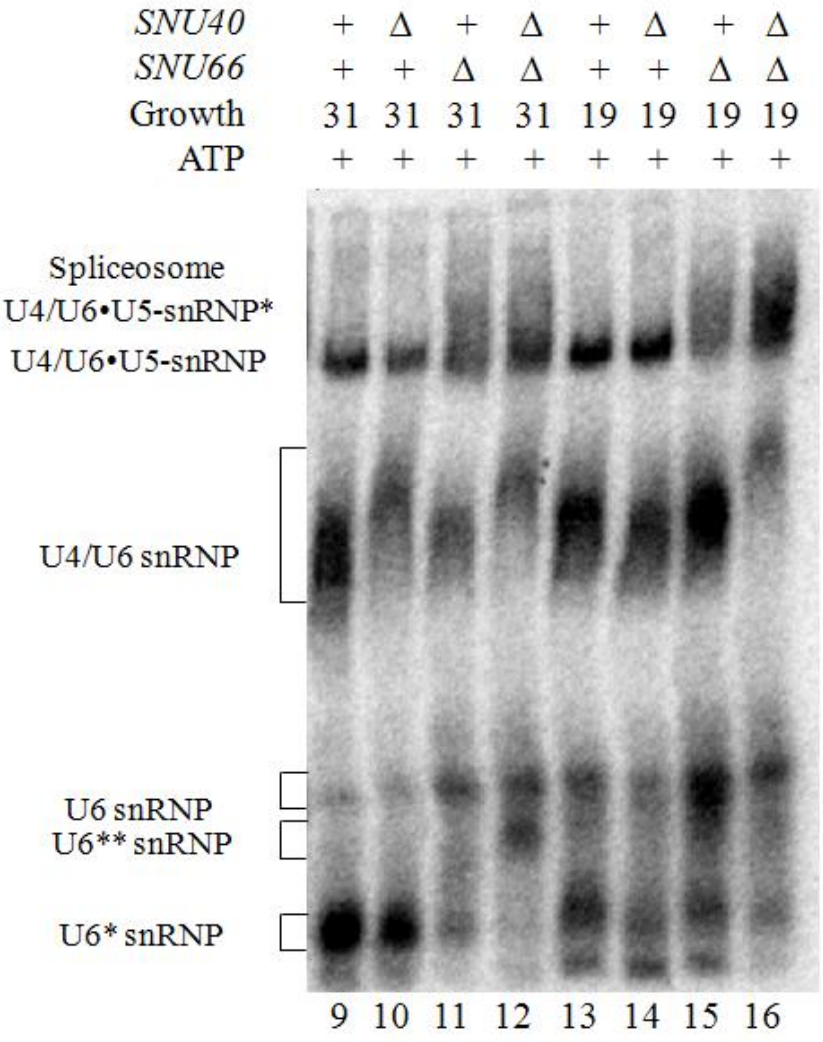


Figure 3.3 - U4/U6•U5 snRNP formation and dissociation in mutant strains

U4/U6•U5 snRNP formation is defective upon removal of SNU66 and restored upon subsequent removal of SNU40. Cells were grown at 31°C and maintained at 31°C or shifted to 19°C until OD (A600) ~ 1.5. Splicing extracts were prepared and incubated at the indicated temperatures prior to electrophoretic separation on native polyacrylamide gels. ATP was added to a final concentration of 2mM to dissociate U4/U6•U5. U4/U6, U6, U6* and spliceosome complexes are indicated.

3.2.4 GTP does not bypass the U4/U6•U5 Disruption Defect

Staley and colleagues have shown that the guanosine nucleotide binding by Snu114p regulates several RNA helicase-dependent steps in splicing (Small, Leggett, Winans, and Jonathan P Staley 2006). We tested whether addition of GTP to the ATP reactions bypassed the ATP resistance and increased U4/U6•U5 snRNP dissociation in Δsnu66 strains. In WT extract,U4/U6•U5 snRNP dissociation was seen upon ATP + GTP treatment, as expected (Figure 3.4 lanes 1-3). While the *Δsnu40* extracts exhibit behavior similar to WT (Figure 3.4 lanes 4-6), U4/U6•U5 snRNP in *Δsnu66* extracts (Figure 3.4 lanes 7-9) and the *Δsnu40/Δsnu66* extracts (Figure 3.4 lanes 10-12) remains resistant to dissociation in the presence of ATP and GTP.

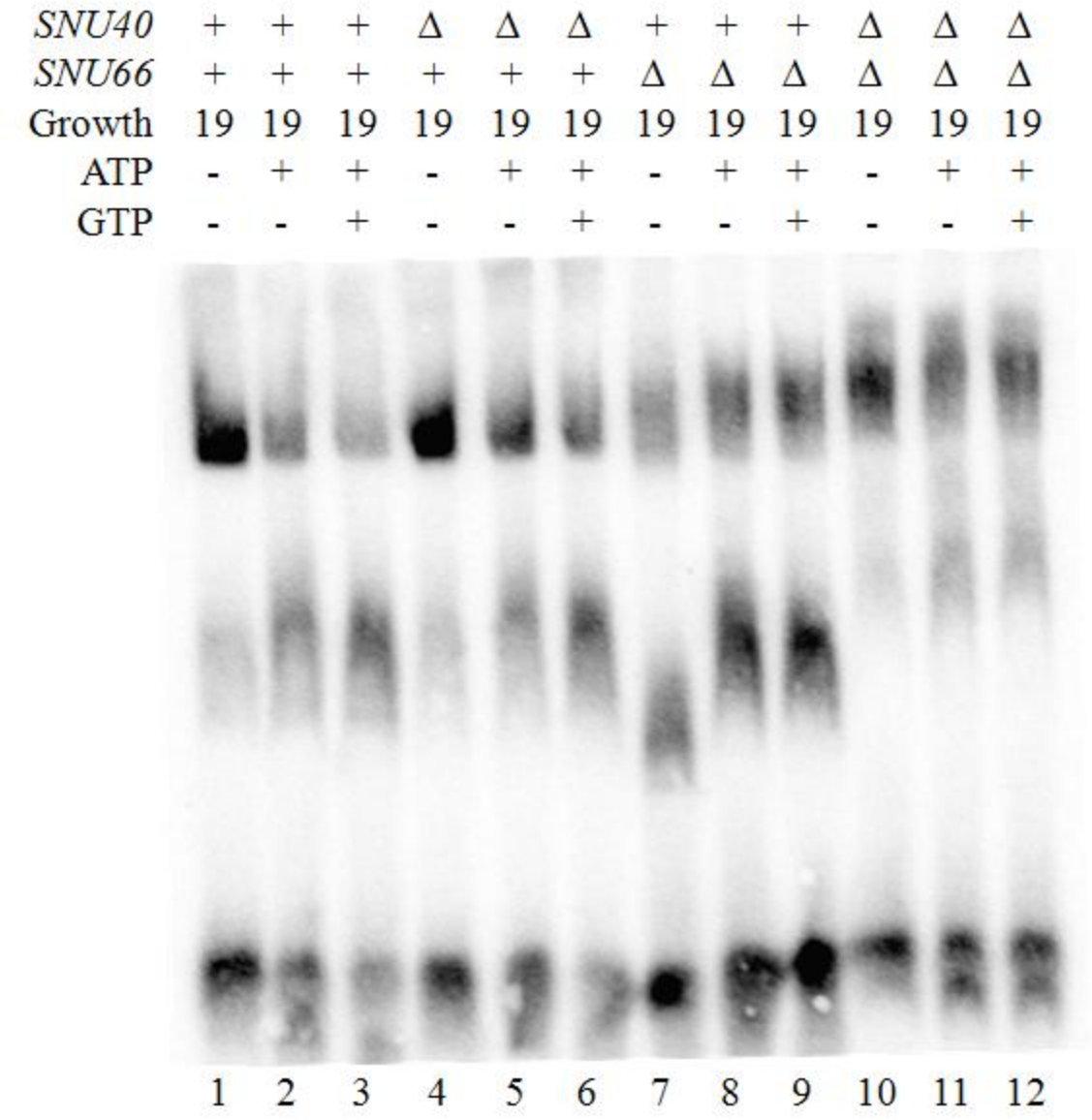


Figure 3.4 – GTP does not enhance the ATP dependent disassociation

Addition of GTP does not enhance the ATP dependent dissociation of U4/U6•U5 snRNP in *Δsnu66* or *Δsnu40/Δsnu66* extracts. Temperature shifted (19°C) extracts were prepared as in Figures 3.3 and 3.4 but supplemented without nucleotide, with ATP or with ATP + GTP. (All ATP and GTP additions resulted in a 2mM final concentration.)

3.2.5 Removal of SNU66 Decreases U4/U6 unwinding

Disruption of WT U4/U6•U5 snRNP upon ATP addition results from unwinding U4/U6 base-pairing, and likely, disruption of protein-protein interactions. To determine whether the U4/U6•U5 snRNP in *Δsnu66* and *Δsnu40/Δsnu66* extracts was incapable of U4/U6 unwinding or if U4/U6 was unwound but remained associated due to inappropriate protein-protein interactions in the mutants, we performed a U4/U6 unwinding assay (Raghunathan and C. Guthrie 1998b). The WT, Δsnu40, Δsnu66 and Δsnu40/Δsnu66 strains were modified to contain a TAP tag (Puig et al. 2001) on the U5 component Brr2p. Brr2p-containing material was affinity captured, washed and incubated in buffer containing 2 mM ATP. Denaturing northern blot analysis of the supernatant and snRNAs remaining on the beads shows that in the WT extract, ATP drives efficient U4/U6 unwinding and release (Figure 3.5 lanes 1-3). The efficiency of this activity is unaffected in Δsnu40 extracts (Figure 3.5 lanes 6-8). In *Δsnu66* extracts, however, decreased unwinding of U4 and U6 was seen (Figure 3 lanes 11-13). The *Δsnu40/Δsnu66* mutant strain similarly exhibited decreased U4/U6 unwinding efficiency (Figure 3, lanes 16-18). Similar to that shown in Figure 2B, treatment of the Brr2p-containing material with ATP + GTP did not enhance the ability to unwind U4/U6 in the complexes (Figure 3.5 lanes 4,5; 9,10; 14,15; 19,20).

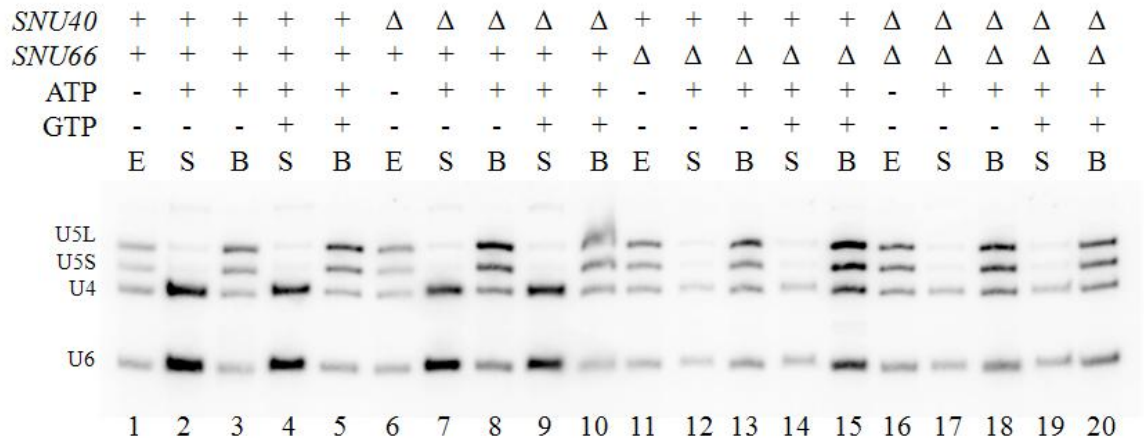


Figure 3.5 – U4/U6 duplex unwinding

U4/U6 duplex unwinding is deficient in Δsnu66 and Δsnu40/Δsnu66 extracts. Brr2p-TAP containing snRNPs were captured on IgG sepharose, washed and incubated with buffer containing ATP or ATP + GTP. Extracted nucleic acids from total Brr2p-TAP snRNP (E), ATP-treated supernatant (S) or beads after ATP treatment and washing (B) were separated on denaturing RNA gels, transferred to nylon membranes and northern blotted with U4, U5 and U6 snRNA probes.

3.2.6 Aberrant Complexes in Δsnu66 and Δsnu40/Δsnu66

The WT and deletion strains were modified with a TAP tag on Prp6p. Extracts from each of these strains, and the Brr2p-TAP strains, were analyzed by native gel analysis and western blotting for the TAP tag (Figure 3.6). Remarkably, WT cells grown at 31°C and extract incubation at 31°C shows very little Prp6p in U4/U6•U5 snRNP (Figure 3.6A lane 1). In each of the mutant strains, there is very little to no Prp6p detected in a U4/U6•U5 snRNP-sized band in extracts prepared from cells grown at 31°C (Figure 3.6A lanes 2-4). Prp6p is seen stably associated with the U4/U6•U5 snRNP in cells shifted to 19°C prior to extract preparation and subsequent incubation at 19°C (Figure 3.6A lanes 5-8) indicating a stabilization of Prp6p at lower temperatures.

Interestingly, in *Δsnu66* cells, much of the Prp6p signal resides in a spliceosome-sized band (Raghunathan and C. Guthrie 1998a) in the 31°C samples, and in an aberrantly migrating U4/U6•U5 snRNP-like band (U4/U6•U5 snRNP*) in the 19°C samples (Figure 3.6A lanes 7,8). Treatment of identical extract samples with 2 mM ATP does not dissociate the spliceosome-sized bands, nor the U4/U6•U5 snRNP* species in Δsnu66 extracts (data not shown). These results were confirmed using the U4/U6•U5 snRNP-associated Brr2p as well (Figure 3.6B) although unlike Prp6p, Brr2p is a stable component of the U4/U6•U5 snRNP at 31°C. These data indicate that in Δsnu66 and *Δsnu40/Δsnu66* cells, the most snRNP-associated Prp6p and Brr2p are found in an aberrantly migrating complex or locked in the spliceosome, however the migration of U4/U6•U5 snRNP* is not detectably different from the migration of U4/U6•U5 snRNP in glycerol gradients (see below). This is true at both permissive and non-permissive temperatures in *Δsnu66* and *Δsnu40/Δsnu66* extracts, although the effect is more severe at low temperatures.

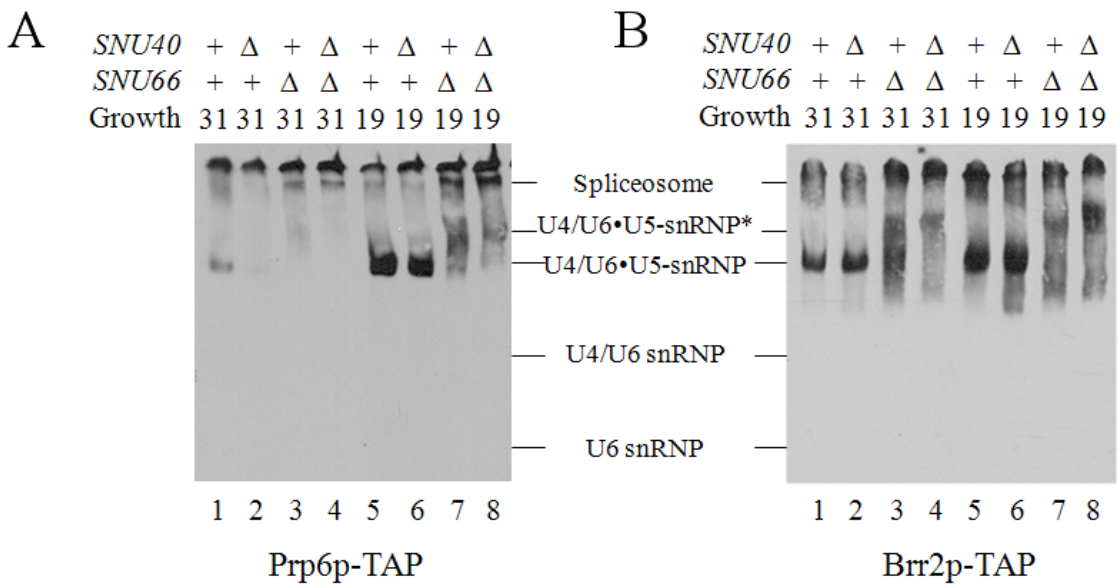


Figure 3.6 - Prp6p and Brr2p are contained in larger complexes

A. Prp6p-TAP extracts with the indicated SNU40 and SNU66 genotypes were separated in non-denaturing polyacrylamide gels, transferred to nitrocellulose membranes and western blotted with PAP. Spliceosome and U4/U6•U5 complexes and the predicted positions of the U4/U6 and U6 snRNPs are indicated.

B. Brr2-TAP extracts with the indicated SNU40 and SNU66 genotypes were analyzed as described in [A].

3.2.7 Transient Association of Prp6p with U5 snRNP

Although Prp6p is not a stable, stoichiometric component of the U5 snRNP in yeast (S. W. Stevens et al. 2001), we hypothesized that if there was collaboration between Snu40p and Snu66p in U4/U6•U5 snRNP assembly, that there may be a transient intermediate in which Prp6p is found with some fraction of the U5 snRNP. To determine whether this was the case, and to determine how the *Δsnu40* and *Δsnu66* deletions might affect this interaction, we purified the Prp6p-containing material by TAP affinity purification from WT, *Δsnu40* and *Δsnu66* strains. TEV protease-eluted material was sedimented through a glycerol velocity gradient and the snRNAs in the fractions were detected by northern blotting in from denaturing PAGE gels (Figure 3.7). When performing similar experiments with Brr2p, the U5 to U4/U6•U5 snRNP ratio is typically 1:1 ((S. W. Stevens et al. 2001); SWS unpublished data), however the ratio of U5 to U4/U6•U5 snRNP in the Prp6p-containing material from the WT strain was very low as determined by phosphorimage quantitation (1:19) (Figure 3.7A). In *Δsnu40* extracts, the proportion was reduced by more than half (1:43) indicating a diminished interaction of Prp6p with the U5 snRNP in the absence of Snu40p (Figure 3.7B). In *Δsnu66* extracts, the proportion of Prp6p-associated U5 to U4/U6•U5 snRNP was increased by more than 2 fold (1:9) suggesting that the transient Prp6p-U5 association was stronger, or lasted longer in the absence of Snu66p (Figure 3.7C).

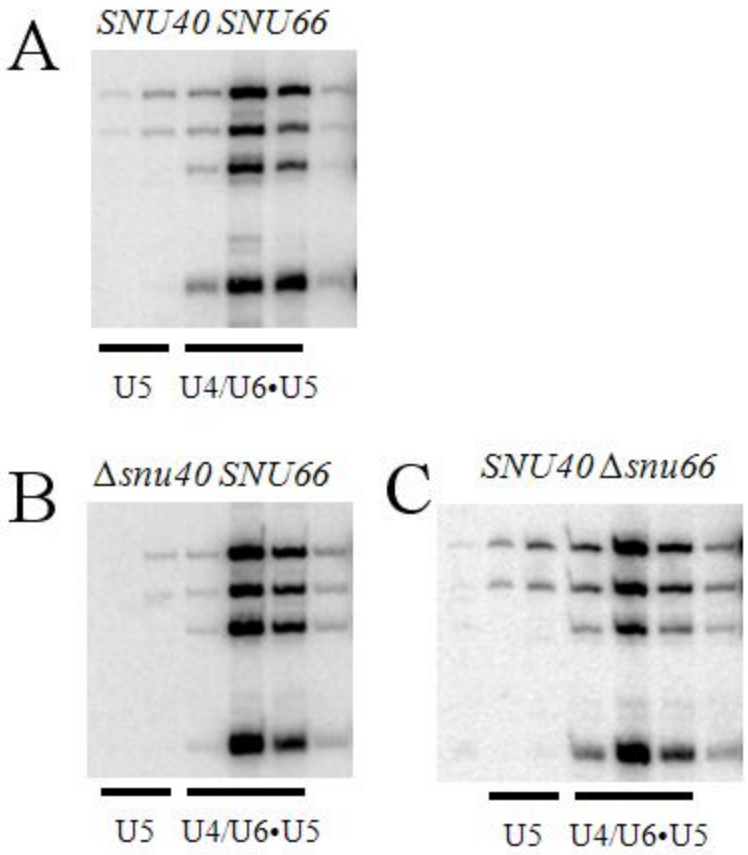


Figure 3.7 – Prp6-TAP immunoprecipitates U5 snRNP

A small portion of the U5 mono-snRNP is affinity-purified using Prp6-TAP. Prp6-TAP affinity-purified material (TEV eluate) was layered onto glycerol velocity gradients. Gradients were fractionated and nucleic acids and proteins were separated by organic extraction. RNA was separated on a urea-PAGE gel. Nucleic acids were transferred to nylon membranes and hybridized with radiolabeled oligonucleotides to detect U4, U5 and U6 snRNAs. Samples from isogenic SNU40/SNU66 [A], *Δsnu40* [B], *Δsnu66* [C] are shown. Locations of the U5 mono-snRNP and U4/U6•U5 snRNP are denoted below.

3.2.8 Temperature Dependent Prp6p Stability in snRNP Complexes

Previous studies of have shown that U4/U6•U5 snRNP in yeast and humans contains Prp6p (S. W. Stevens et al. 2001)(Scott W. Stevens and John Abelson 1999)(A. Gottschalk et al. 1999)(Behrens and R. Lührmann 1991). We note that in WT extracts incubated at 31°C, Prp6p becomes de-stabilized from the U4/U6•U5 snRNP, and that incubation at 19°C stabilized Prp6p association with U4/U6•U5 snRNP (Figure 3.6 lanes 1,5). Although previous studies have shown the human (Evgeny M. Makarov, Olga V. Makarova, Tilmann Achsel, and Reinhard Luhrmann 2000) and yeast (Abovich, Legrain, and Rosbash 1990) orthologs of Prp6p are present in the U5 and U4/U6 snRNPs respectively, our data are in conflict as there is no Prp6p-TAP signal in a U4/U6-sized complex in any strain at either temperature in the native gels (Figure 3.6A) nor was it a stoichiometric component in the stable U5 snRNP from yeast (S. W. Stevens et al. 2001). Examining the association of the TAP-tagged Prp6p at 19°C via native gel showed Prp6p to be stably associated with the U4/U6•U5 snRNP as well as spliceosome in WT and *Δsnu40* strains. To address the association of Prp6p with different complexes in these extracts in a manner not dependent on native gel electrophoresis, we sedimented whole-cell extracts through glycerol gradients to separate free Prp6p from complexes (Figure 3.8). In WT, *Δsnu40* and *Δsnu66* extracts, Prp6p was primarily associated with U4/U6•U5 snRNP-sized complexes, but to a lesser extent, trailing off into upper fractions (Figure 3.8A, 3.8B, 3.8C). Whether this trailing effect is due to association of Prp6p with different complexes or dissociation during the gradient analysis is difficult to determine, but the latter is more likely given the results presented in Figure 3.6A. In the *Δsnu40/Δsnu66* extracts, however, Prp6p was found throughout the gradient indicating a significant destabilization from splicing complexes (Figure 3.8D).

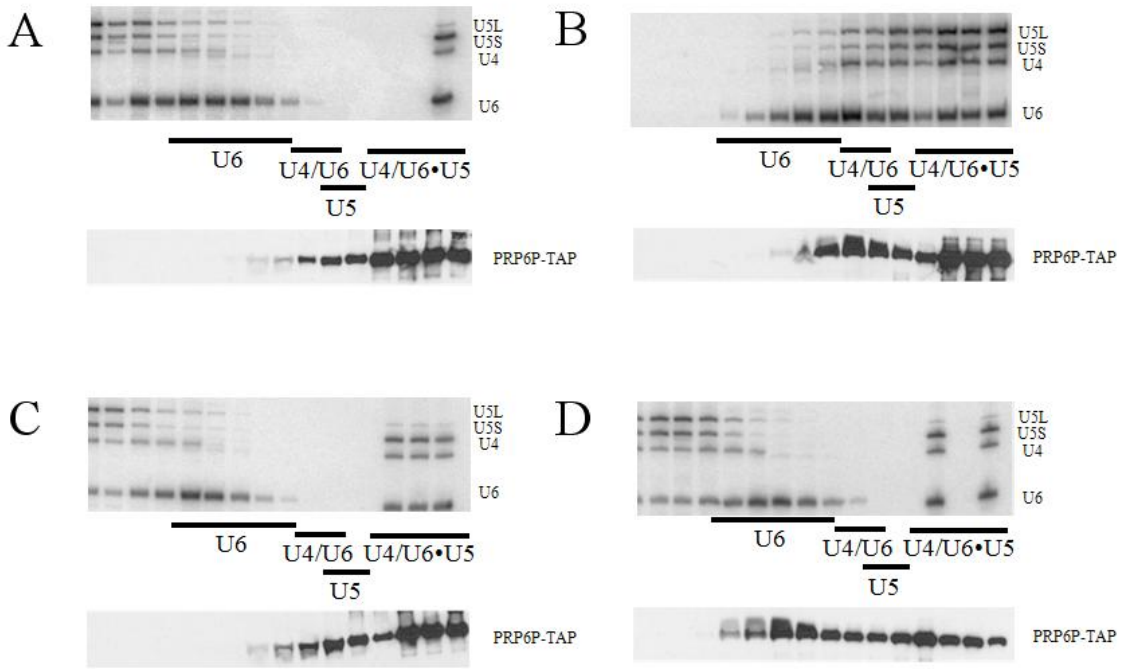


Figure 3.8 – Prp6p is destabilized in a *ΔSNU40+ΔSNU66* mutant

Extract separated on a gradient from Prp6p-TAP tagged strains. [A] contains both *SNU40* and *SNU66*, while [B] lacks *SNU40*. [C] lacks *SNU66* and [D] lacks both *SNU40* and *SNU66*.

3.3 DISCUSSION

Details involving how cells assemble and maintain proper levels of the spliceosomal snRNPs remain largely unknown. Additionally, the presence of myriad forms of some snRNPs (such as U5;((S. W. Stevens et al. 2001)(Scott W. Stevens and John Abelson 1999)(A. Gottschalk et al. 1999)(Behrens and R. Lührmann 1991)(O.V. Makarova, E.M. Makarov, H. Urlaub, C.L. Will, Gentzel, Wilm, and R. Lührmann 2004)(A. Gottschalk, B. Kastner, R. Luhrmann, and P. Fabrizio 2001)(Bach, Winkelmann, and R. Lührmann 1989))) raises several questions as to how eukaryotic cells functionally discriminate between them and properly maintain directionality in their formation and usage. In WT cells, Snu40p is found in the U5 snRNP but not the U4/U6•U5 snRNP and Snu66p is found only in the U4/U6•U5 snRNP. Both Snu40p and Snu66p have demonstrated interactions with Prp6p. Here we have characterized the role of Snu40p and Snu66p in U4/U6•U5 snRNP formation and function and the relevance of their interactions with Prp6p.

While deletion of *SNU40* does not result in a U4/U6•U5 snRNP assembly defect, the deletion of *SNU66* does block formation of this particle at low temperatures (Figure 2A). The *Δsnu66* strain accumulates the U4/U6 particle under these conditions indicating a defect in U4/U6•U5 snRNP assembly. This is in contrast to the human ortholog which can be immunodepleted from salt-treated extracts with little effect on U4/U6•U5 snRNP assembly (O.V. Makarova, E.M. Makarov, and R. Lührmann 2001).

The association of Prp6p prior to entry into the yeast U4/U6•U5 snRNP has been a subject of some debate, with previous data including it in the U4/U6 snRNP (Abovich, Legrain, and Rosbash 1990) while data regarding the human protein places it in the U5 snRNP (Evgeny M. Makarov, Olga V. Makarova, Tilmann Achsel, and Reinhard

Luhrmann 2000). Our data indicate that Prp6p is only a stable binding partner in the U4/U6•U5 snRNP and larger complexes (Figure 3.6A), however it can be seen as a transient interacting partner with a small portion of U5 snRNP (Figure 3.7). An explanation for this discrepancy is that Prp6p-associated yeast U4/U6•U5 snRNP was dissociated at the elevated salt levels used which left Prp6p with the U4/U6 portion of the disassembled U4/U6•U5 snRNP, while the human homologue has a stronger association with the U5 components. Nonetheless, given our data and the mammalian interaction data between Snu40p and Prp6p orthologs (Laggerbauer et al. 2005), yeast Prp6p forms a transient intermediate with the U5 snRNP through its interaction with Snu40p prior to U4/U6•U5 snRNP formation.

Based on the previously characterized interactions and data presented here, we propose a model for the functions of Snu40p, Snu66p, and Prp6p in U4/U6•U5 snRNP assembly (Illustrations 3.1-3.4). In WT cells, Prp6p interacts transiently with Snu40p in the U5 snRNP. As the U4/U6 snRNP joins the U5 snRNP to create the U4/U6•U5 snRNP, Prp6p is handed off to the U4/U6•U5 snRNP specific Snu66p with concomitant Snu40p release (Illustration 3.4). In the *Δsnu40* strain,Prp6p is able to properly load into the U4/U6•U5 snRNP via its interaction with Snu66p ((S.B. Liu, Rauhut, H.P. Vornlocher, and R. Lührmann 2006); Illustration 3.1). In the Δsnu66 strain, the process proves to be more problematic, as Snu40p and Prp6p remain intertwined. This interaction blocks U4/U6•U5 snRNP formation and leads to increased U4/U6 snRNP levels as well as a >2-fold enrichment of Prp6p-associated U5 intermediate (Illustration 3.2). Furthermore, the U4/U6•U5 snRNP that forms in these circumstances is deficient in Brr2p unwinding activity, likely due to some regulatory function of Snu66p upon Brr2p (S.B. Liu, Rauhut, H.P. Vornlocher, and R. Lührmann 2006)(van Nues and Jean D. Beggs 2001). The Δsnu66 U4/U6•U5 snRNP formation block is relieved by subsequent

deletion of SNU40 (Illustration 3.3), indicating that these proteins serve to function via a common interacting partner, Prp6p. Despite efficient U4/U6•U5 snRNP formation at permissive temperatures, the *Δsnu40/Δsnu66* strain is dead at 16°C (Figure 3.1). The cold-sensitive phenotype is likely the result of a U4/U6•U5 snRNP particle that is unable to efficiently unwind U4 and U6 snRNAs. Analysis of Brr2p unwinding activity (Figure 3.5) confirms the reduced unwinding activity of Brr2p in the Δsnu66 strain as well as the double deletion strain. Furthermore, added ATP (Figure 3.5 lanes 18-19) or ATP and GTP (Figure 3.5 lanes 19-20) in *Δsnu40/Δsnu66* extracts does not dissociate U4/U6•U5 snRNP levels relative to the non-ATP supplemented extract. The reduced U4/U6 unwinding activity should inevitably lead to a buildup of the spliceosome just prior to activation, likely at the A2-1 step (S.C. Cheng and J. Abelson 1987). Indeed, increased spliceosome levels are seen in strains lacking Snu66p (Figure 3.2, Figure 3.6A).

An additional mechanism by which the A2-1 complex may stall is through malformation of the spliceosome due to a misplaced or unstable Prp6p. Assuming that the purpose of Snu40p and Snu66p are to ensure proper loading of Prp6p into the U4/U6•U5 snRNP, it follows that deletion of these chaperones results in instability and likely, malfunction under non-optimal conditions. The behavior of Prp6p in a glycerol gradient (Figure 3.8D) confirms that it is most destabilized from the U4/U6•U5 snRNP in *Δsnu40/Δsnu66* extract. It is likely these additive effects are responsible for elevated levels of spliceosomal complexes in the double deletion in comparison to the *Δsnu66* strain (Figure 3.6A lanes 7, 8). Others have noted that non-essential gene products functioning in essential cellular processes often have regulatory roles which only affect cell function or cell growth under certain conditions (Fang, Rocha, and Danchin 2005). Snu40p and Snu66p appear to be functioning in a similar manner in spliceosome assembly. Their regulation of U4/U6•U5 snRNP assembly is not essential in rich medium

at optimal growth temperatures. However when cells are stressed by decreasing the growth temperature, the optimal assembly of the U4/U6•U5 snRNP is disrupted in the absence of Snu66p.

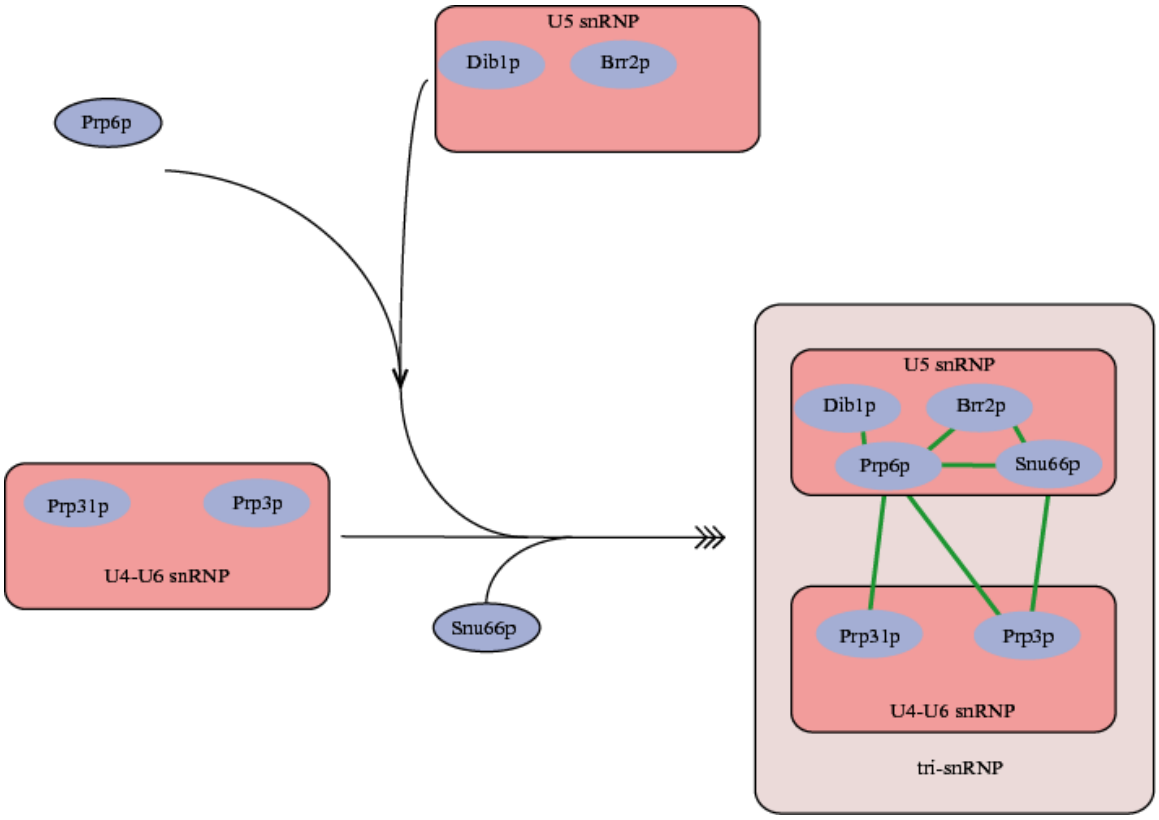


Illustration 3.1 – Tri-snRNP loads Prp6p in a *ΔSNU40* strain

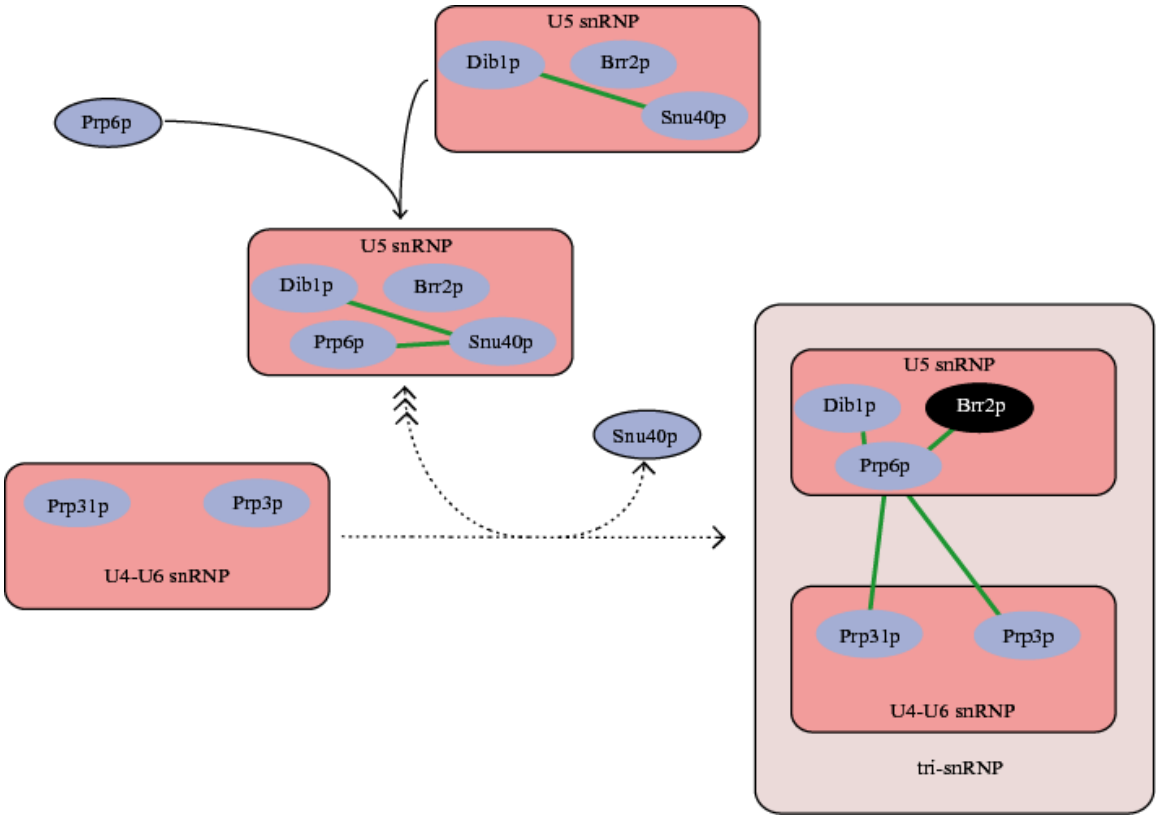


Illustration 3.2 – Snu40p retains Prp6p without Snu66p

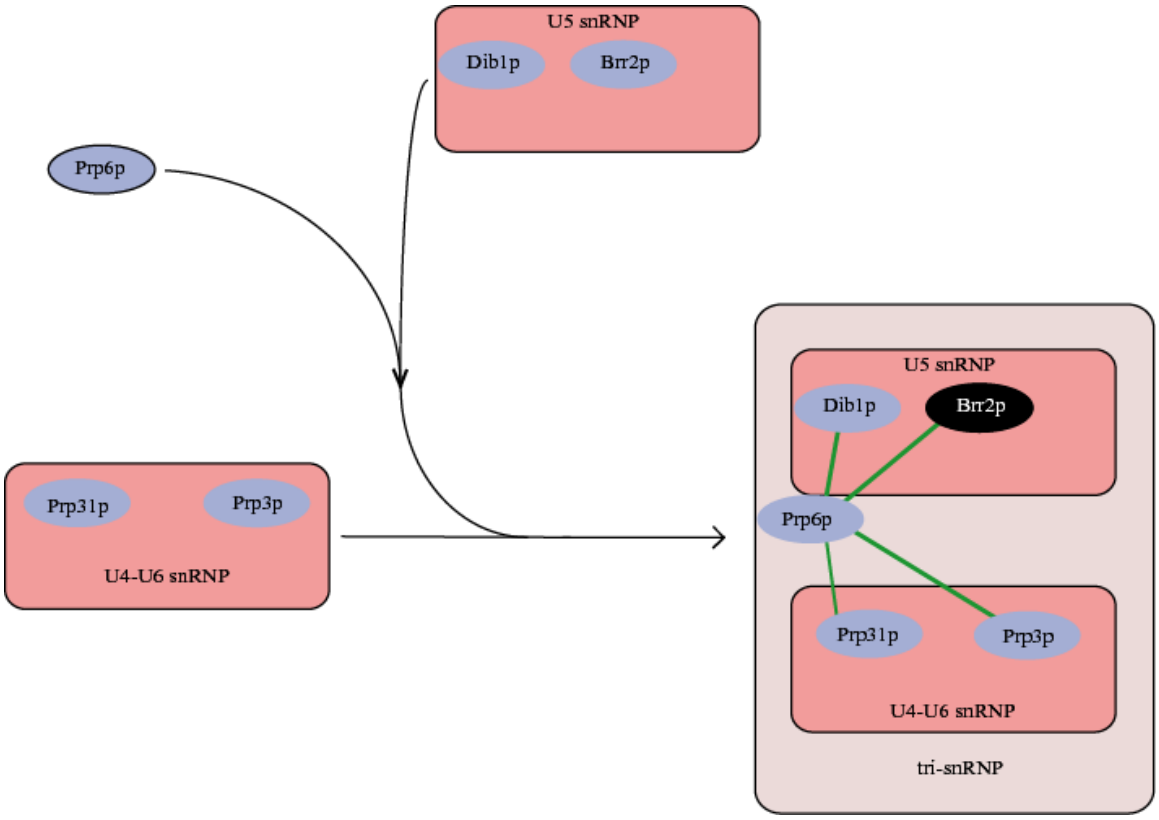


Illustration 3.3 – Prp6p is destabilized from tri-snRNP in *ΔSNU40* + *ΔSNU66* mutants

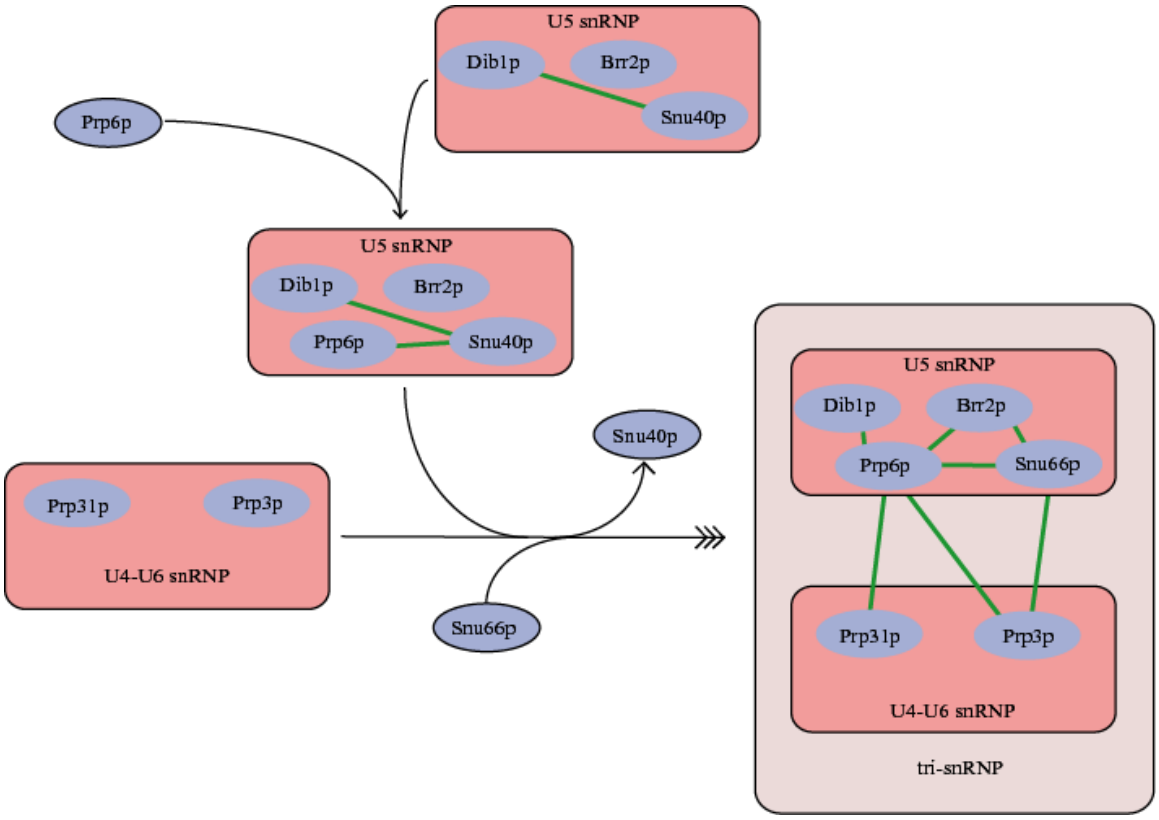


Illustration 3.4 – A model for tri-snRNP formation

Keeping this new model of tri-snRNP formation in mind it is informative to revisit the previously published two hybrid and biochemical interaction data that was presented in the first chapter. Illustration 3.5 arranges this data with this new structure in mind, and finds that the previous interactions overlap superbly with our new functional data.

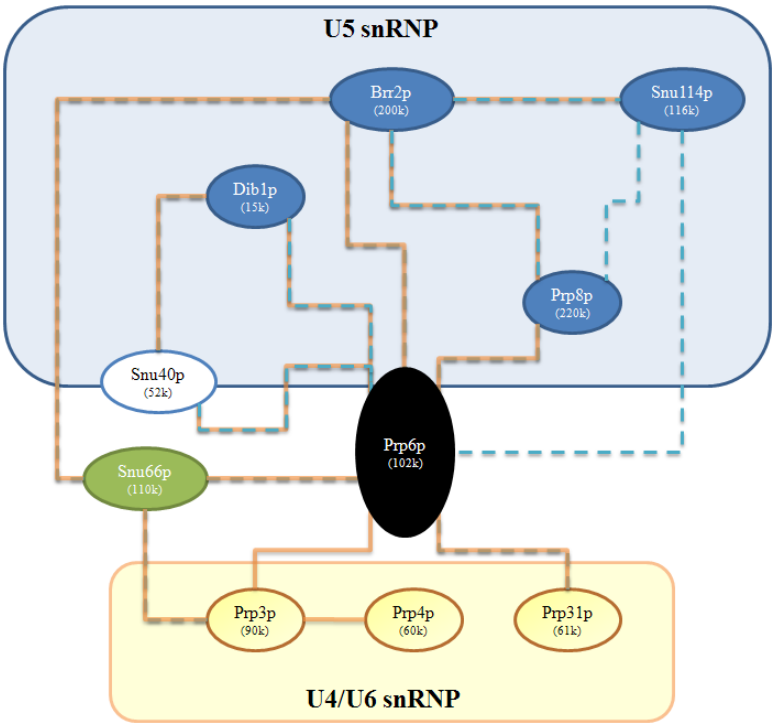


Illustration 3.5 – Revisited protein-protein interactions leading to the tri-snRNP

Saccharomyces cerevisiae tri-snRNP interactions are combined with data from their human homologs (subtitled in parenthesis). Dashed lines indicate yeast two-hybrid interactions, while solid lines indicate co-immunopurification data. Proteins in blue are found in both U5 as well as the tri-snRNP. Snu40p, in white, is U5 specific. Proteins found in both U4/U6 and tri-snRNP are colored yellow. Snu66p, in green, is tri-snRNP specific. Lastly, Prp6p/102k is found in tri-snRNP in both species but has been reported to be a U5 specific protein in humans. Note that this figure represents protein interactions only. Snu40p and Snu66p are never simultaneously located in the same particle. Figure based on data from (van Nues and Jean D. Beggs 2001)(Olga V Makarova, Evgeny M Makarov, Sunbin Liu, Hans-Peter Vornlocher, and Reinhard Lührmann 2002)(S.B. Liu, Rauhut, H.P. Vornlocher, and R. Lührmann 2006)(Gonzalez-Santos et al. 2002)..

CHAPTER 4: PURIFICATION OF A FUNCTIONAL PRE-CATALYTIC SPLICEOSOME

4.1 INTRODUCTION

In the process of spliceosome assembly and activation several rearrangements occur, largely by the action of RNA helicases (Tanner and P. Linder 2001). While this process was explored in depth in the first chapter, a brief review follows. The process (Illustration 4.1) begins with the base pairing of the U1 snRNA to the 5' splice site forming the commitment complex (Krainer 1997). Following this event, branchpoint binding proteins engage their target (Rutz and B. Seraphin 1999) with the U2 snRNP soon following. At this point, the B-complex or Pre-Spliceosome is formed (S.C. Cheng and J. Abelson 1987). The addition of a preformed tri-snRNP and other protein factors create the precatalytic spliceosome (A2-1 complex) (S.C. Cheng and J. Abelson 1987). After a series of ATP-driven base pairing rearrangements, the U1 and U4 snRNPs are released and the activated spliceosome (A1 complex) is formed. These steps see the Prp28p dependent unwinding of U1 from the 5' splice site (J P Staley and C Guthrie 1999) as well as the Brr2p driven unwinding of U4 from U6 (Xu, Nouraini, Field, Tang, and Friesen 1996)(Lauber et al. 1996)(D. H. Kim and Rossi 1999). These actions free both the intron of 5' splice site and U6 snRNA for future base pairing. At this time the U5 snRNA concurrently addresses the exon at the 5' splice site. The completion of these events result in the formation of the A1 complex of the spliceosome. The next rearrangement of the spliceosome results from the actions of Prp2p, which moves the branchpoint adenosine into close proximity of the 5' splice site (Silverman et al. 2004).

The steps that follow complete the splicing cycle through two transesterification reactions, release of the mature mRNA, and recycling of all RNP components.

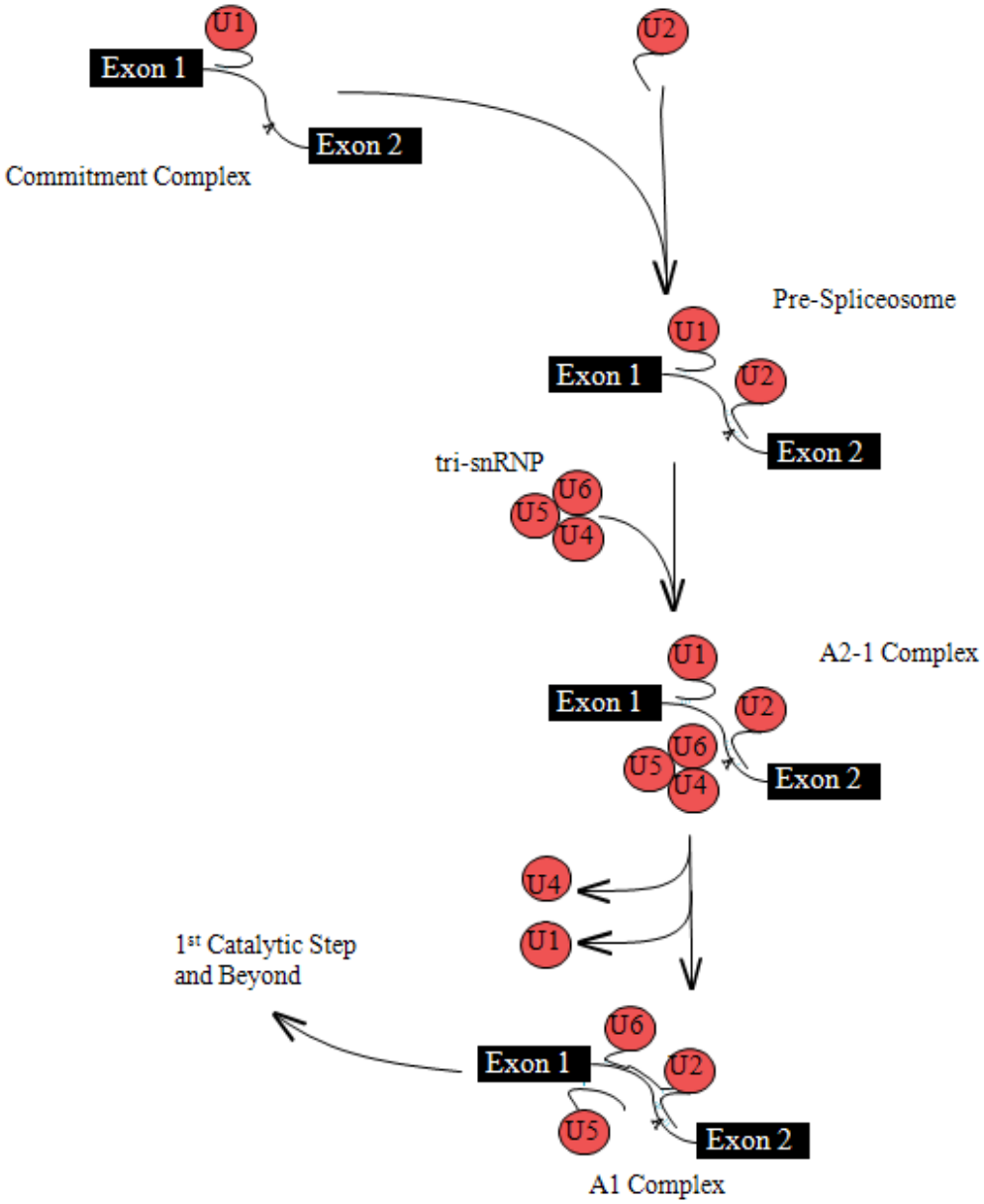


Illustration 4.1 - Rearrangements prior to the 1st catalytic step of splicing

4.1.1 Formation of the Activated Spliceosome

The reductionist approach used in the past has resulted in a modest understanding of the composition and movements of the component snRNPs that comprise the spliceosome. Complexes like the U4/U6 di-snRNP and the U4/U6·U5 tri-snRNP are relatively well characterized (Nottrott, H. Urlaub, and R. Lührmann 2002)(Raghunathan and C. Guthrie 1998a)(Scott W. Stevens and John Abelson 1999)(S. W. Stevens et al. 2001). Nevertheless, much remains unknown about the larger spliceosome complexes like the fully assembled, precatalytic A2-1 complex. In addition to the substrate pre-mRNA, this complex contains all five snRNAs (S.C. Cheng and J. Abelson 1987). Like the B complex before it, A2-1 includes a pre-mRNA that is base paired with U1 at the 5' splice site while U2 remains base paired at the branchpoint sequence. The addition of tri-snRNP to the complex sets in motion the process that leads to the activated spliceosome. Namely, the U6 snRNA must basepair with both the 5' splice site as well as the U2 snRNA (D.A. Wassarman and J.A. Steitz 1993). Additionally, the Lsm protein ring must be removed from U6 (Chan, Der-I Kao, Wei-Yu Tsai, and Soo-Chen Cheng 2003). This requires not only the Prp28p-dependant release of the U1 snRNA from its position on the mRNA (J P Staley and C Guthrie 1999), but also the unwinding of the previous U4/U6 snRNA pairing. Brr2p, in concert with Snu114p, is responsible for the unwinding of this extensively basepaired U4/U6 duplex (Raghunathan and C. Guthrie 1998b)(Small, Leggett, Winans, and Jonathan P Staley 2006). Once free of U4, U6 is then able to base-pair with the substrate while U4 is released from the spliceosome and fed into a snRNP recycling pathway. After U6 is base-paired with the 5' splice site, the remodeling continues via addition of a number of protein factors termed the Nineteen Complex (NTC) (Chan and Soo-Chen Cheng 2005).

Prp19p is known to be associated with seven other proteins (Cef1p, Isy1p, Ntc20p, Snt309p, Syf1p, Syf2p, Syf3p) to form the NTC (S.C. Cheng, W.Y. Tarn, T.Y. Tsao, and J. Abelson 1993)(W.Y. Tsai et al. 1999)(W.Y. Tarn, Hsu, K.T. Huang, H.R. Chen, H.Y. Kao, K.R. Lee, and S.C. Cheng 1994)(Hau-Ren Chen, Shr-Peng Jan, Twee Y. Tsao, Sheu, Josette Banroques, and Soo-Chen Cheng 1998)(C.H. Chen, W.Y. Tsai, H.R. Chen, C.H. Wang, and S.C. Cheng 2001)(Chun-Hong Chen et al. 2002). It has been reported that the NTC is not found in the pre-spliceosome (W.Y. Tarn, K.R. Lee, and S.C. Cheng 1993). Nevertheless, the complex plays a pivotal role in the adaptation of U6 during the transition to the active complex (Chan and Soo-Chen Cheng 2005). Within these parameters, the exact sequence of NTC arrival to the spliceosome remains a matter of debate. Even beyond this, however, the presence of the NTC in splicing complexes represents a broader debate between the stepwise model of assembly and the penta-snRNP. The previous characterization of the penta-snRNP complex found all eight core NTC components to be present (Scott W. Stevens et al. 2002). This data may indicate that it is the true functional, on pathway intermediate. Thus, under the preformed spliceosome hypothesis, *in vitro* purifications using harsh salt conditions have been characterizing a subset of the factors in the *in vivo* pre-A2-1 complex. Therefore, omission of the NTC from these datasets may be biochemical artifact.

What is known about the NTC is that it is required for stable association of U5 and U6 with the spliceosome after the exit of U1 and U4 (Chan, Der-I Kao, Wei-Yu Tsai, and Soo-Chen Cheng 2003). This stabilization is concurrent with a shift in U6 basepairing and release of the Lsm protein ring (Chan, Der-I Kao, Wei-Yu Tsai, and Soo-Chen Cheng 2003). The next step in the process is a Prp2p-dependent rearrangement of the spliceosome in preparation the for the first transesterification reaction (J.H. Chen and R.J. Lin 1990)(King and J.D. Beggs 1990). Noteworthy in this step is the requirement of

Spp2p for the binding of a free Prp2p to the spliceosome (Roy, K. Kim, Maddock, Anthony, and Jr 1995). After the Prp2p rearrangement, the NTC-related factor Yju2 is thought to promote the first transesterification reaction of splicing (Yen-Chi Liu, Hsin-Chou Chen, Nan-Ying Wu, and Soo-Chen Cheng 2007).

4.1.2 The Assembled Spliceosome in Yeast versus Mammals

While a large body of data has been generated from yeast, a number of biochemical studies have recently emerged from mammalian extracts. In so doing, a number of differences have been identified in the splicing machinery. Some of these basics were reviewed in chapter one of this dissertation. Some are reviewed below.

Perhaps the largest difference between the mammalian splicing apparatus and that of *Saccharomyces cerevisiae* is in the lack of the alternative AT-AC spliceosome machinery in yeast (Woan-Yuh Tarn and Joan A. Steitz 1997). This system utilizes pre-mRNA with non-canonical splice sites and relies on a different set of snRNPs with the exception of the shared U5 (Claudia Schneider, Cindy L Will, Olga V Makarova, Evgeny M Makarov, and Reinhard Lührmann 2002). This point is worth noting on a technical level, as a pulldown against any U5 component will result in a heterogeneous population in the mammalian system. There are no such complications in budding yeast.

Nonetheless, many differences also exist between the canonical yeast and mammalian spliceosomes. Purified mammalian spliceosome have contained a number of proteins that have no known yeast counterpart (Appendix I). The reverse of this situation presents itself when analyzing yeast complexes. Generally speaking, these differences may represent species-specific function or simply result from uncharacterized trans-acting factors.

4.1.3 tri-snRNP Formation and the Impact on the Spliceosome

It is in the re-formation of the tri-snRNP that the catalytic core of the spliceosome is assembled. This particle is the combination of three snRNAs and 28 proteins (S. W. Stevens et al. 2001). After the splicing cycle is complete, Prp24p aids the reannealing of free U4 to U6 snRNA to form the U4/U6 di-snRNP (Raghunathan and C. Guthrie 1998a). This particle associates with the U5 snRNP to form the U4/U6·U5 tri-snRNP. The basepairing that brought together U4 and U6 continues to pair these RNAs in the tri-snRNP. Unlike U4 and U6, however, U5 appears to be associated with the particle exclusively through protein-protein interactions. Previously, Prp6p was identified as a possible candidate (Galisson and Legrain 1993) for bridging U5 to U4/U6. Mutations in *PRP6* reduced tri-snRNP formation while levels of the precursor U5 and U4/U6 snRNPs remained unaffected. We have recently confirmed this role of Prp6p through the *SNU40* and *SNU66* deletions documented in Chapter 3. To summarize, Prp6p was shown to interact transiently with the U5 snRNP. This U5 interaction takes place via Snu40p, which ultimately coordinates the handoff of Prp6p to Snu66p during tri-snRNP formation. A deletion of both *SNU40* and *SNU66* results in the destabilization of Prp6p from the tri-snRNP and ultimately a defect in U4/U6 unwinding. The reduced ability of Brr2 to unwind U4 and U6 results in the observed cold sensitive phenotype.

4.2 RESULTS

4.2.1 Double Deletion Strains Accumulate PRP6 in a large particle.

In chapter three, data was exhibited that the deletion of *SNU40* and *SNU66* resulted in decreased Brr2p unwinding of U4 from U6. Given this fact, one could hypothesize that a stalled pre-catalytic spliceosome accumulates under restrictive conditions. Indeed, another look at the native gel analysis of Prp6p containing particles revealed that the

deletion of *SNU66* resulted in the buildup of large complexes in whole cell extracts (Figure 4.1, a subset of the earlier Figure 3.6). To review, cell extract was separated by native gel, transferred to a membrane and analyzed by western blot utilizing an antibody against the TAP tag affixed to Prp6p. In the strain containing both *SNU40* and *SNU66* the vast majority of Prp6 is contained in the U4/U6·U5 complex at 19°C (Figure 4.1, lane 1). This pattern was also seen in the strain lacking *SNU40* (Figure 4.1 lane 2). Nonetheless, when *SNU66* is deleted from the strain, Prp6p is destabilized from the tri-snRNP. Further, Prp6p can be detected in a spliceosome sized complex. (Figure 4.1 lane 3). When both *SNU40* and *SNU66* are deleted, the resulting extract contains Prp6p almost exclusively in the spliceosome sized particle (Figure 4.1 lane 4).

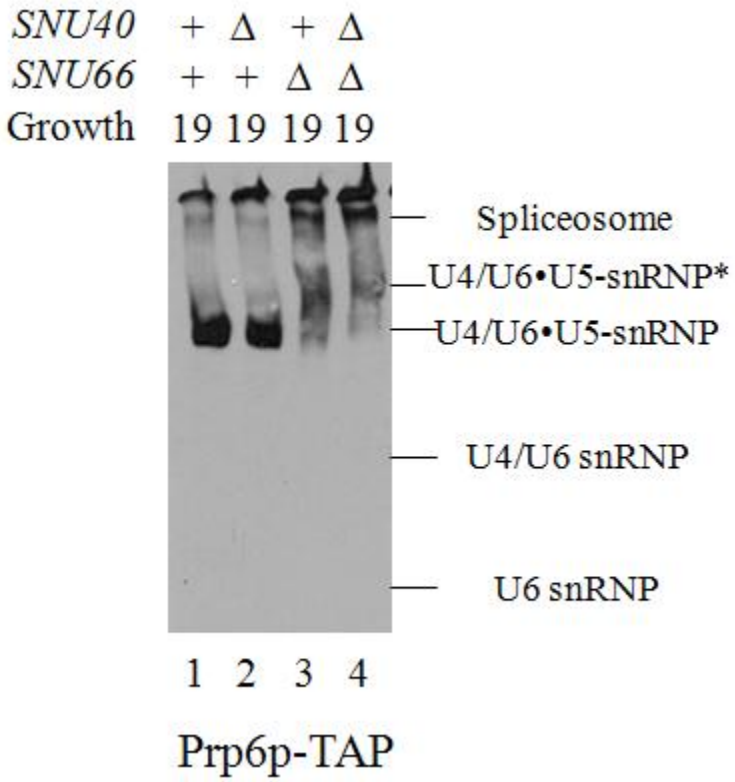


Figure 4.1 - Δ*snu40*/Δ*snu66* extracts contain Prp6p in large complexes

Prp6p-TAP tagged strains created from the indicated *SNU40* and *SNU66* genotypes were used in the creation of whole cell extract. These extracts were separated on a non-denaturing polyacrylamide gel and transferred to a nitrocellulose membrane. This membrane was then western blotted with a PAP antibody. The positions of spliceosome and U4/U6•U5 complexes are indicated while the predicted positions of the U4/U6 and U6 snRNPs are inferred.

4.2.2 Purification of a large PRP6 containing complex

The buildup of Prp6p in a spliceosome-sized complex provides an interesting target for purification and analysis. To this end, after harvesting cells previously temperature shifted to 19°C, we utilized the TAP tag on Prp6p for affinity purification. After cleavage from the IgG-sepharose column with Tobacco Etch Virus (TEV) protease, a glycerol gradient was used to separate the Prp6p-associated particles. After fractionation of the gradient, the samples were phenol-chloroform extracted and the nucleic acid-containing aqueous layer was separated from the protein containing organic layer. The subsequent aqueous fractions were then precipitated and analyzed by denaturing gel northern blot analysis with ^{32}P-labeled oligonucleotide probes against U1, U2, U4, U5, and U6 snRNAs (Figure 4.2). Fractions 7-9 (~25S) contain tri-snRNP, and all five snRNAs were found in fractions 15-17. This position in the gradient corresponds to a size of ~55S. It is also noteworthy that all five were found in a roughly equal stoichiometry.

The protein content from each fraction in the glycerol gradient was also analyzed. After acetone precipitation, protein pellets were resuspended in LDS loading buffer and resolved on a denaturing Bis-Tris gel. A number of protein bands can be observed in the lanes determined to contain U4/U6·U5 (Figure 4.3 lanes 7-9) as well as all five snRNAs (Figure 4.3 lanes 15-17).

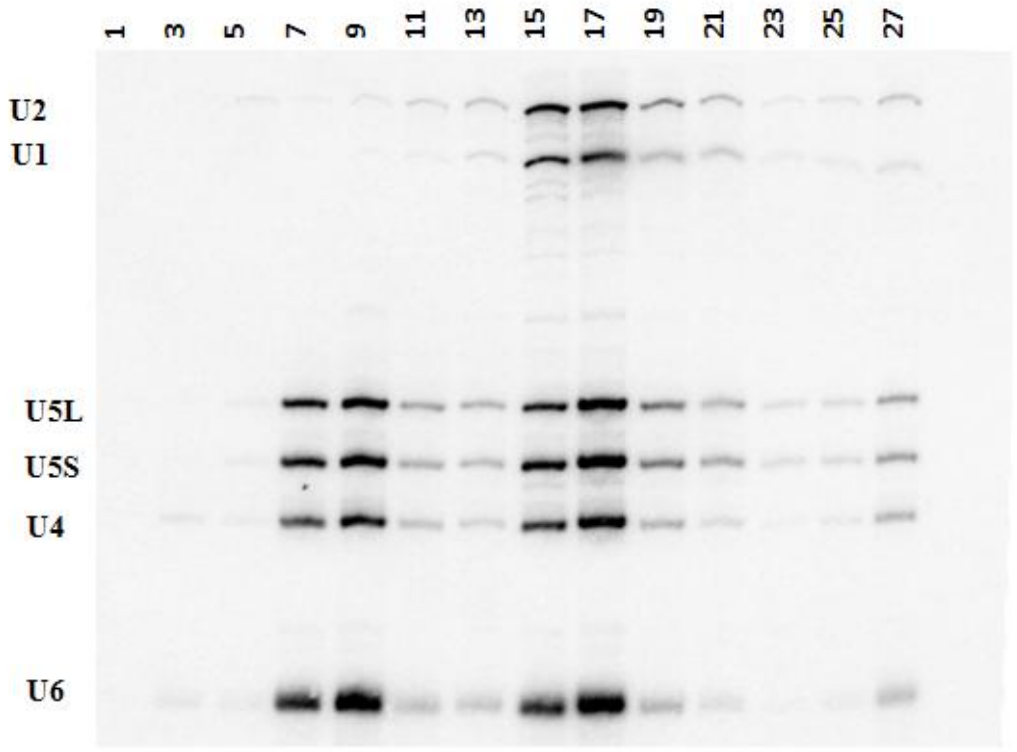


Figure 4.2 – All five snRNAs are found in a 55S complex

All five snRNAs are found in a ~55S sized complex in a temperature shifted Δ*snu40/*Δ*snu66* extract following PRP6-TAP affinity purification. Following dialysis to 50mM NaCl, whole cell extract was affinity purified by IgG-Sepharose and TEV elution. Eluate was placed on a 10%-30% glycerol gradient, fractionated, and the RNA was precipitated. Odd numbered fractions were run on a denaturing gel. Northern blot analysis using probes against all five snRNAs indicate the presence of all of these snRNAs in fractions 15-19.

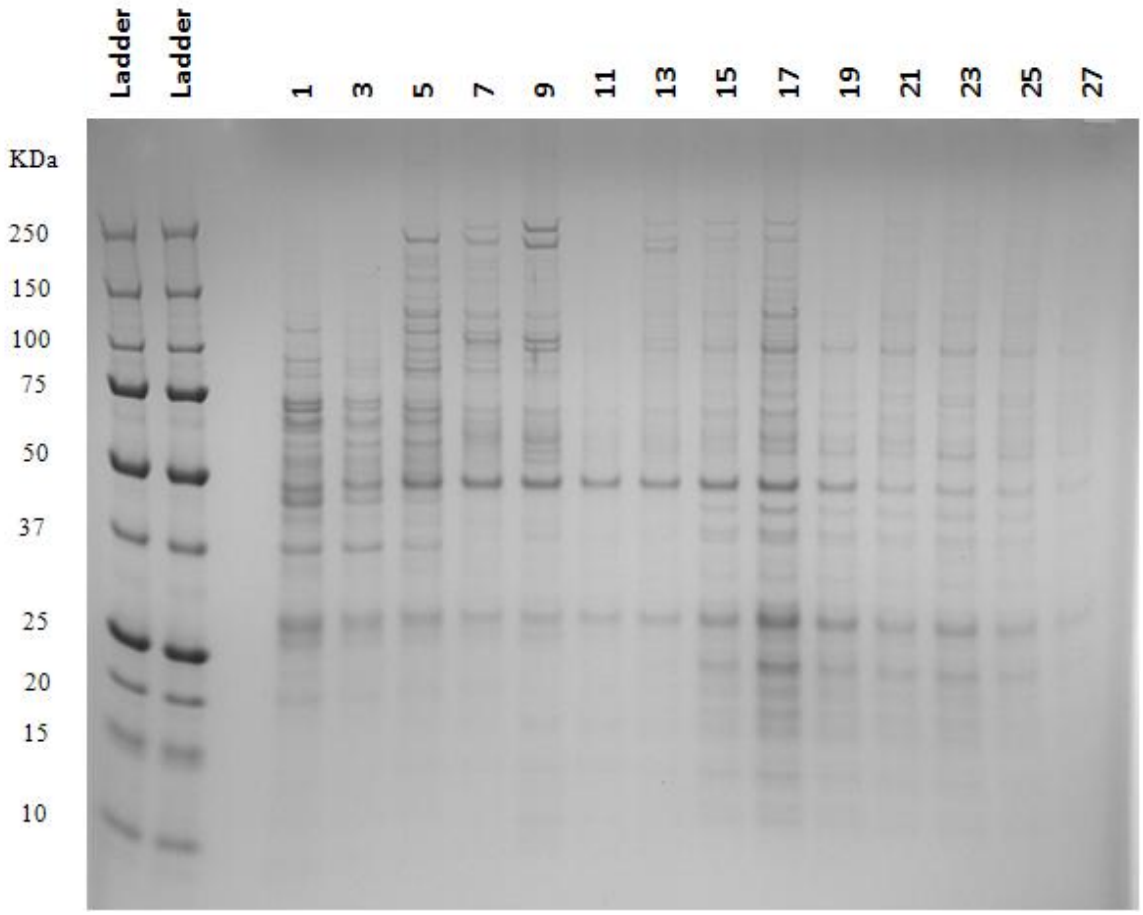


Figure 4.3 – A protein complex is identified to migrate at 55S

Proteins were precipitated from the fractions collected in Figure 4.2. A 4%-12% BIS-TRIS gel separated proteins, which were then visualized by Coomassie stain.

4.2.3 U6 is exclusively base paired with U4.

4.2.3.1 Native RNA Analysis

Previous work has shown that U6 undergoes numerous rearrangements throughout the splicing cycle. The basepairing of U4/U6 in the context of all five snRNAs is one of the defining characteristics of the pre-catalytic A2-1 spliceosome. To examine this basepairing status we employed two different methods. The first was to examine RNA from the purified particle under native conditions. The second examined psoralen crosslinked RNA. RNA analysis was ultimately performed by northern blotting with probes for U2, U4, and U6. Pooled glycerol gradient material (Fractions 15-19) identical to that prepared in the previous section was used.

The native gel analysis required use of both a native control as well as one prepared under denaturing conditions. The first control lane (Figure 4.4) indicates the size of a denatured, and therefore un-basepaired, mono-U6 snRNA. The second control lane, prepared under native conditions, exhibits the presence of both basepaired and mono-U6 snRNA. In the third lane, the RNA from purified spliceosome indicates that fully 100% of the U6 snRNA is base-paired with a partner snRNA. While this assay identified the base pairing status of U6, the snRNA that formed the other half of that pairing remained unknown. As described above, U6 is found base-paired with U4 in the snRNPs and in the precatalytic spliceosome, and with U2 in the catalytic spliceosome.

.

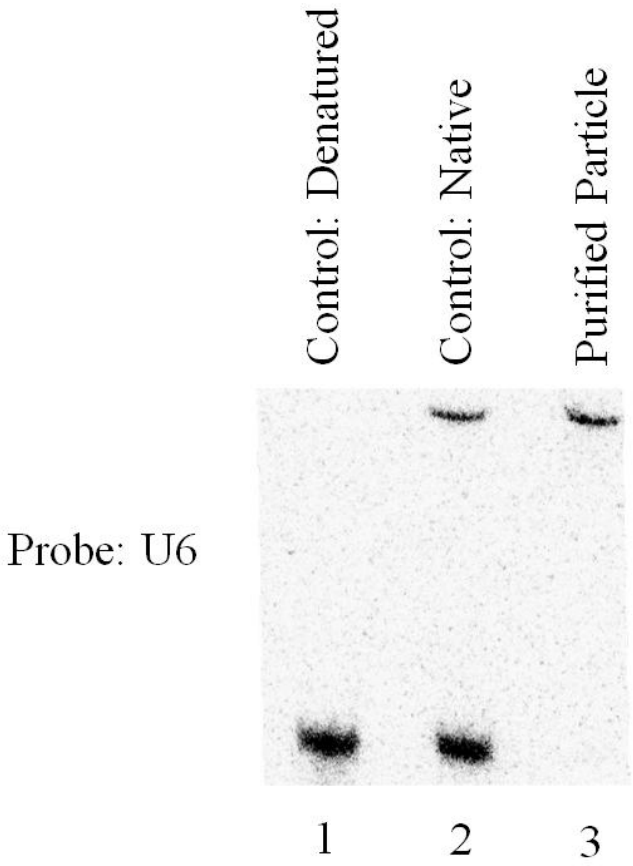


Figure 4.4 - 100% of the U6 snRNA is base paired in purified particle

The first two lanes are controls that identify the migration rate of a base paired or denatured U6 snRNA. The first lane consists of whole cell RNA prepared under denaturing conditions, while the second lane contains whole cell RNA prepared under denaturing conditions. The third lane contains 55S glycerol gradient material after phenol chloroform extraction and a gel load under native conditions. After transfer to a membrane, the sample was probed for the U6 snRNA. These results indicate that the U6 snRNA contained in the particle is exclusively base paired with a single species.

4.2.3.2 Psoralen Crosslinking of RNA

To identify any possible U2/U6 or U4/U6 base pairing, the same glycerol gradient material used previously in section 4.2.3.1 was employed. The RNAs were crosslinked by addition of psoralen and application of UV light (K.M. Wassarman and J.A. Steitz 1993). After phenol chloroform extraction and ethanol precipitation, the RNA was resuspended, split, and run in three identical lanes under denaturing conditions. This gel was then cut into slices and transferred onto three different membranes which were then probed for either the U2, U4, or U6 snRNAs (Figure 4.5). While U2 was bound to a large complex contained near the well, U4 and U6 co-migrated to identical positions on the gel. Confirming the results in section 4.2.3.1, no free mono-U6 was seen.

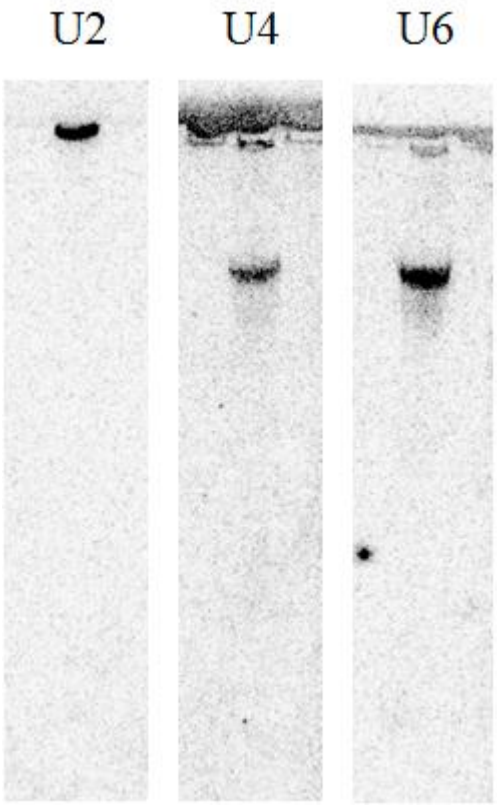


Figure 4.5 - Crosslinked RNA reveal U4-U6 basepairing

Psoralen crosslinking of purified particle reveals U4-U6 base pairing and not U2-U6. 55S sized glycerol gradient material was crosslinked with Psoralen and run on a denaturing gel. The gel was split and transferred to multiple membranes that were in turn individually probed for U2, U4, and U6.

4.2.4 Pre-mRNA Status in the Active Complex

Previous work (Scott W. Stevens et al. 2002) indicated that all five snRNAs can exist in the absence of a substrate pre-mRNA. Thus, in order to confirm or reject this complex as the A2-1 pre-catalytic spliceosome, the presence or absence of pre-mRNA was tested. Given the extremely low quantity of any single species of pre-mRNA, selecting an assay with adequate sensitivity was a concern. In response, RT-PCR was used to confirm the status of three separate transcripts: RPS4A, RPS10A, and RPL16A. To prepare the RNA for the reaction, proteinase-K treatment was utilized to rid the post-glycerol gradient particle of all protein components. The remaining RNA was then phenol–chloroform extracted followed by ethanol precipitation. Figure 4.6 contains complete data on RPS4A while RPS10A and RPL16A data is summarized in Figure 4.7

The results (Figure 4.6) of the RPS4 RT-PCR reveal that both pre-mRNA and mature mRNA are present. Utilizing primers on both exons of RPS4 results in the identification of full pre-mRNA and as well as the shorter spliced mature mRNA (Figure 4.6 Lane 1). A second primer set, directed against the intron and second exon (Figure 4.6 Lane 2) also resulted in the appropriate sequence length thus confirming the presence of unspliced and/or intermediate transcript. Later sequencing of these DNA bands confirmed the presumed identities. It is noteworthy that a significant amount of mature mRNA is carried in the purified particle, presumably as an external contaminant (see discussion for elaboration). In any event, the presence of pre-mRNA within the context the spliceosome defines the presence of an A2-1 complex.

We then tested the particle for catalytic activity. Splicing reactions were set up in the presence of either wildtype cellular extract (Figure 4.6 Lane 3-4) or RNA-free micrococcal nuclease treated wildtype extract (Figure 4.6 Lane 7-8). Micrococcal

nuclease treatment was executed as described earlier in section 2.6.16 and later in 4.2.5.1. While the introduction of the extract provides factors required for splicing, the introduction of exogenous RNA introduces the possibility that pre-mRNA is spliced by these external players. Additionally, the introduction of additional pre-mRNA or mRNA species could complicate the analysis of the assay results. To prevent any of these issues extract was nuclease treated prior to introduction to the assay.

Controls for these extracts were run to probe our assay's sensitivity to pre-mRNA and mRNA introduced. Extract (Lane 5-6) and nuclease treated extract (Lane 9-10) was tested in the absence of the purified particle. A lack of detectable pre-mRNA and mRNA was observed in both cases.

Overall, the results from this semi-quantitative RT-PCR were plotted for each species (pre-mRNA, intron, and mRNA) as a percentage of total sample RNA (Figure 4.6 bottom). Note that the 'intron' primer set anneals to the pre-mRNA as well as intermediate species and therefore represents a mix of both. The results show that the RPS4A pre-mRNA is reduced with the addition of wildtype cellular extract. The magnitude of this reduction is moderated in the presence of nuclease treated extract. While corresponding increases in the levels of mature mRNA are observed, the levels of the 'intron' species seem steady or to rise throughout. Given the decrease in the full length pre-mRNA species (Exon1-Exon2 Large), the lack of a corresponding decrease in the Intron-Exon2 species indicates the presence of a partial block in the second step of splicing.

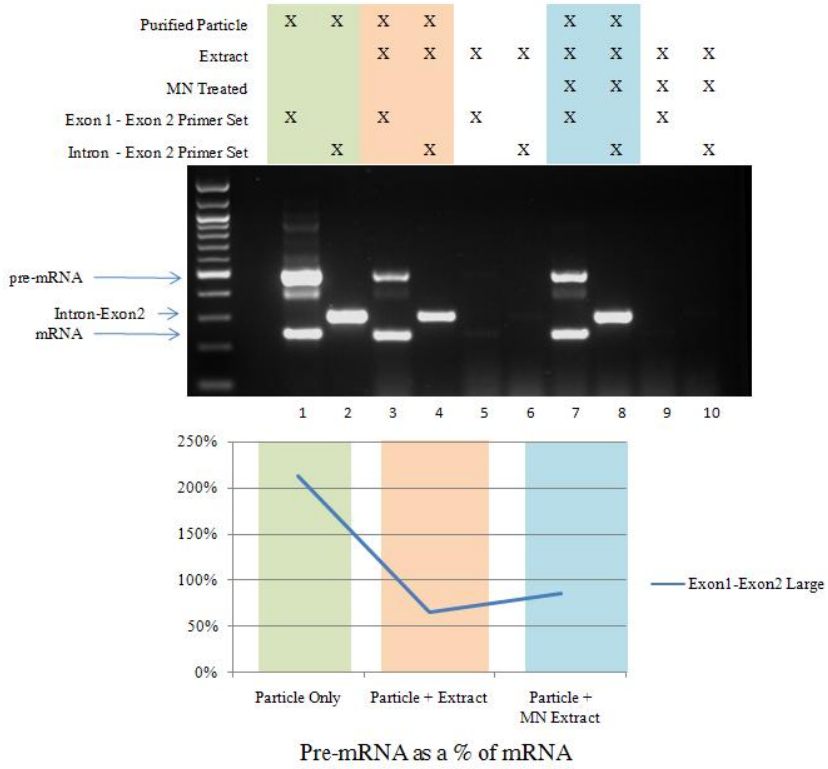


Figure 4.6 - RT-PCR of RPS4 confirms the presence of mRNA

RT-PCR analysis of particle exhibits the presence of both pre-mRNA and mature mRNA. 55S sized glycerol gradient material was proteinase K treated, phenol-chloroform extracted, and ethanol precipitated. Any RNA present was used as a template in an RT-PCR reaction. The reaction product was analyzed by agarose gel. The first experimental lane contains the results using a primer set against both exons in the respective gene. Thus, the larger band represents unspliced pre-mRNA while the lower molecular weight band represents mature mRNA. The second experimental lane contains data resulting from a primer set directed against the intron as well as 3' exon. Thus, the presence of a single band confirms the presence of a mix of pre-mRNA and blocked lariat intermediate.

The analysis of the other two transcripts, RPS10A and RPL16A, resulted in the same general pattern seen for RPS4A in Figure 4.6. For simplicity, analysis of all three transcripts is presented in Figure 4.7. These plots consider mRNA content within the particle itself along with the changes introduced by the nuclease treated extract. In all cases, the band intensity corresponding to the pre-mRNA falls relative to the mature mRNA. Nevertheless, the primer set annealing to the intron and second exon either gains in intensity (RPS4A) or remains roughly static (RPS10A, RPL16A, data not shown). Overall, the behavior of the intron-exon two primer set RT-PCR results indicate a potential partial second step splicing block.

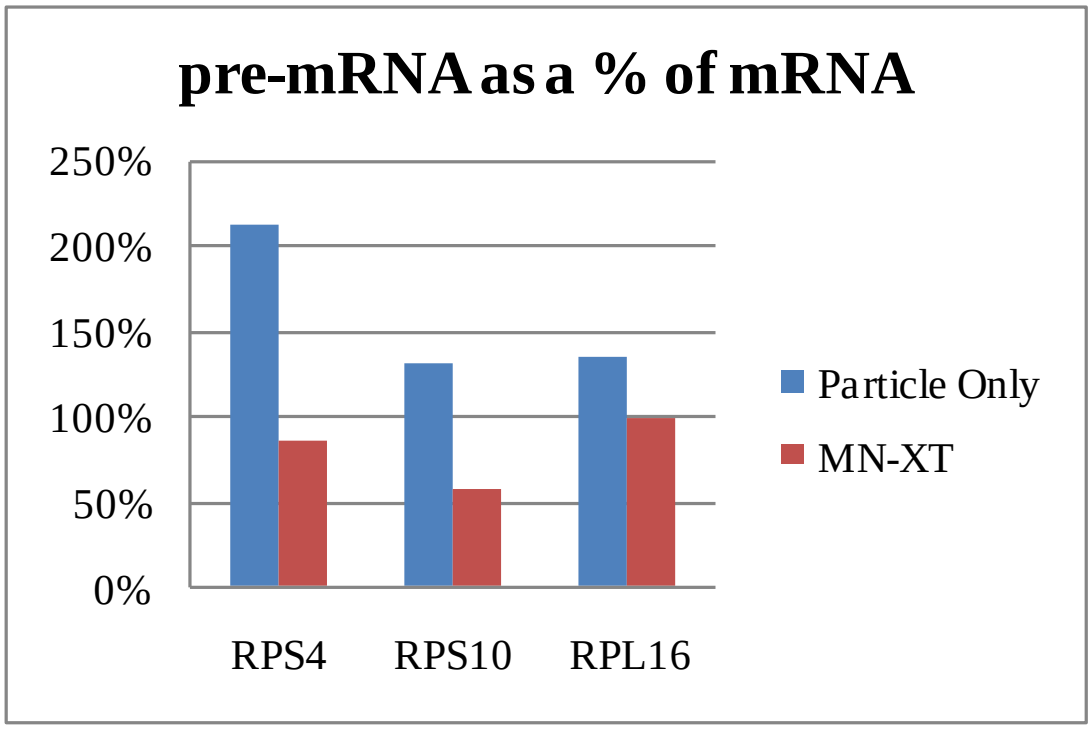


Figure 4.7 – RPS4A, RPS10A, and RPL16 splicing efficiency

4.2.5 Recycling of the A2-1 Complex

4.2.5.1 Micrococcal Nuclease Treatment

In order to assess the vitality of the purified A2-1 complex through multiple rounds of activity, we tested the particle's ability to splice an exogenous mRNA species. The capability of the particle to restore splicing activity to an extract depleted of RNA via micrococcal nuclease indicates that the A2-1 complex is viable for successful recovery and reuse.

We prepared a splicing-competent extract using methods (Ansari and Schwer 1995) similar to that employed earlier in affinity purification. This extract was then treated as has been described elsewhere (Yean and R.J. Lin 1991) with micrococcal nuclease to degrade all RNA at hand. The nuclease activity was later quenched through chelation of calcium ions by addition of EGTA. Northern blot analysis confirmed the ablation of all five snRNAs (Figure 4.8, lane 2).

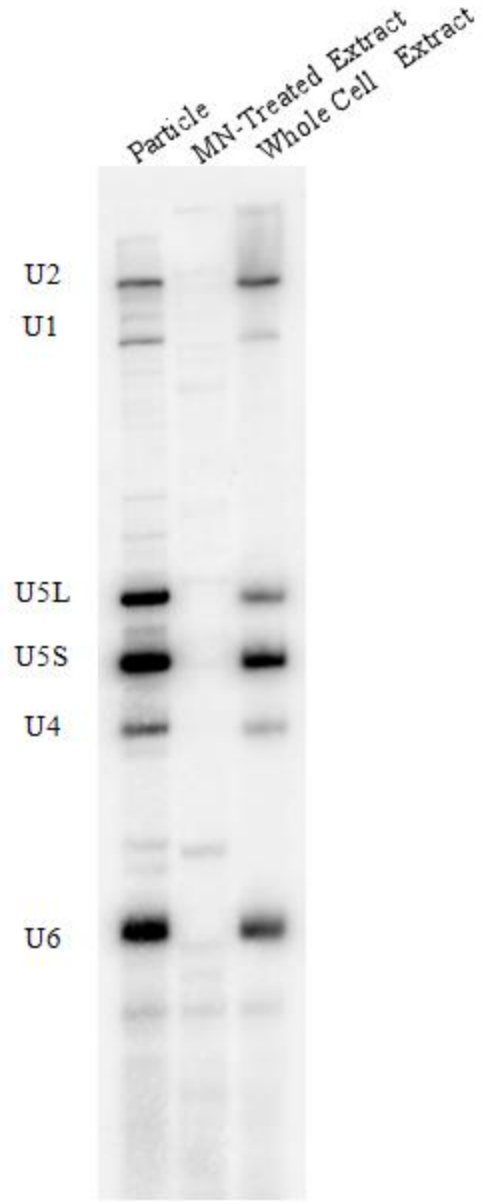


Figure 4.8 - Nuclease degradation of snRNAs

Micrococcal nuclease treatment results in obliterated snRNAs in whole cell extract. Whole cell extract was treated with micrococcal nuclease to degrade any RNA present. Northern blot analysis confirms the successful degradation of all five snRNAs.

4.2.5.2 Splicing Reaction

Upon EGTA chelation of the nuclease reaction, this extract was used in a series of splicing reactions (R.J. Lin, Newman, S.C. Cheng, and J. Abelson 1985). The intent of these reactions was to detect the splicing of an *in vitro* transcribed and radiolabeled pre-mRNA. Negative controls consisted of pre-mRNA only (Figure 4.9, lane 1) and pre-mRNA added to the extract prior to nuclease treatment (Figure 4.9, lane 2). The latter of these controls was intended as an additional method of assessing the success of micrococcal nuclease digestion of the RNAs present. Positive controls included a non-nuclease treated extract (Figure 4.9, lane 3) as well as a non-nuclease treated extract plus purified particle. Among the reactions that used nuclease treated extract, lane 8 served as the negative splicing control as it contained no A2-1 complex. Lanes 5-7 contained increasing amounts of purified A2-1 particle along with nuclease treated extract. (1µl particle in lane 5, 2µl in lane 6, and 4µl in lane 7).

Increasing amounts of purified A2-1 complex resulted in increasing levels of splicing while the control reaction exclusively containing nuclease treated extract exhibited no activity (Lane 8). When analyzed by quantitative software analysis, it can be seen that increasing the volume of purified particle in a splicing reaction from 1µl to 4µl increased levels of spliced mRNA from 17% to 22%, respectively (Figure 4.10). These data can be benchmarked against a non-nuclease treated control reaction (Figure 4.9, Lane 3) which shows 53% of the total RNA being processed to mRNA (Figure 4.10). The addition of 4µl particle (Figure 4.9, Lane 4) to this control has negligible effect with 49% mRNA produced (Figure 4.10). In contrast to the splicing activity of the *in vivo* assembled spliceosome, the recycled splicing complex shows no second step block activity in spite of utilizing identically purified snRNAs.

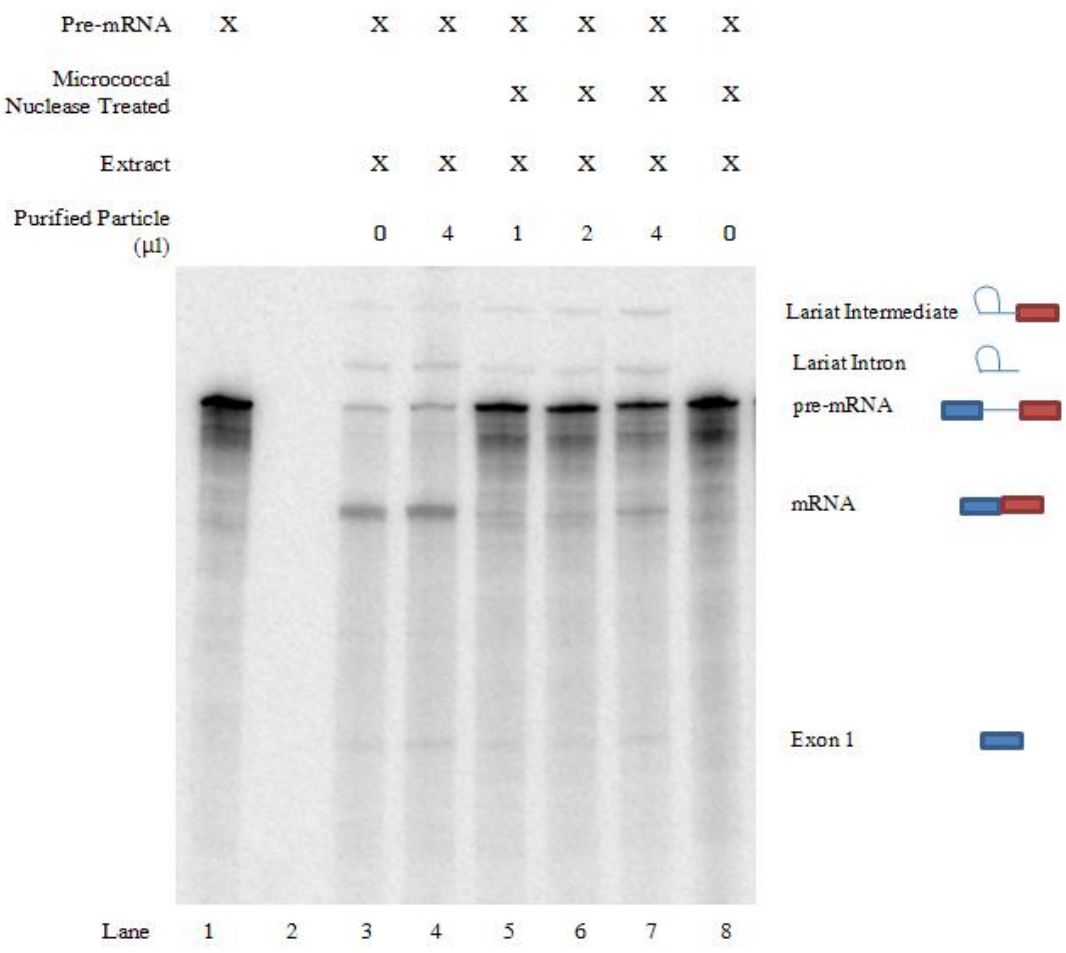
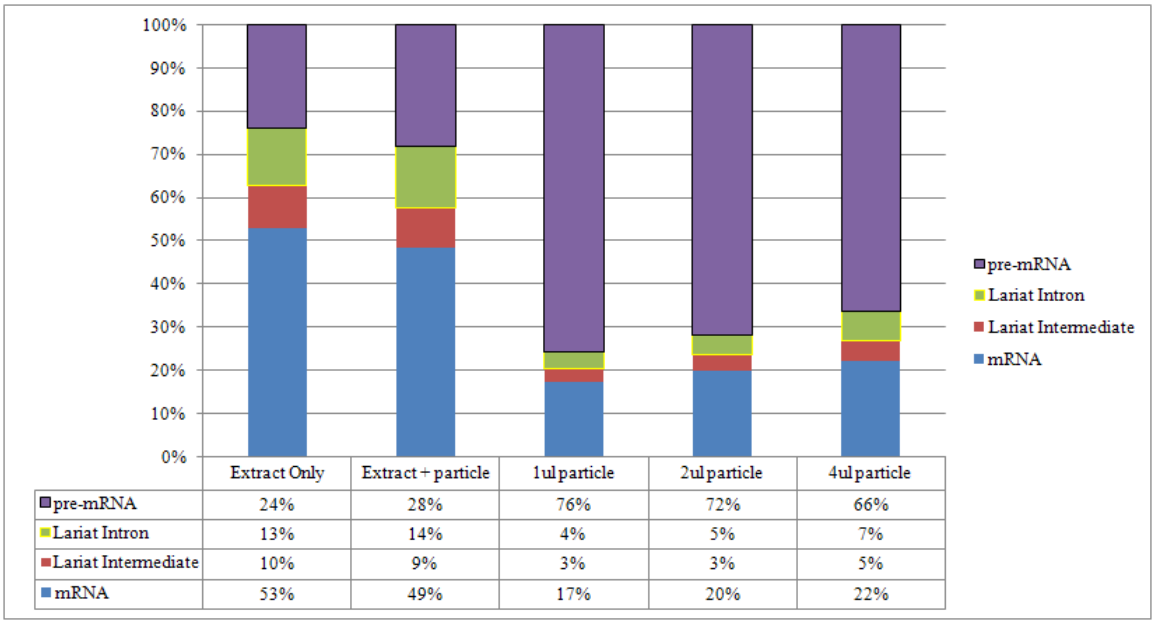


Figure 4.9 – A2-1 complex snRNA recycling

Splicing reactions exhibit catalytic activity of the purified A2-1 complex. Splicing reactions were carried out in the presence of wildtype whole cell extract (lane 1), wildtype extract plus 4µl purified A2-1 particle (lane 2), and increasing volumes of A2-1 complex in micrococcal nuclease treated extract (lanes 3-5). Lane 6 demonstrates the lack of catalytic activity when micrococcal nuclease treated extract is used in the absence of purified A2-1 complex.



	Extract Only	Extract + particle	1ul particle	2ul particle	4ul particle
pre-mRNA	24%	28%	76%	72%	66%
Lariat Intron	13%	14%	4%	5%	7%
Lariat Intermediate	10%	9%	3%	3%	5%
mRNA	53%	49%	17%	20%	22%

Figure 4.10 - Increasing particle results in increased mRNA splicing

Quantitation of bands in Figure 4.9 Lanes 3-7. Exon 1 bands resulting from a 2nd step block were not detectable in amounts sufficient for quantitation.

4.2.6 Three Step Purification Reduces Contamination

The final step in the project led us to submit the peak protein fraction (Figure 4.3, Fraction 17) to mass spectrometric analysis (as described in section 2.6.8). The results (data not shown) of this analysis unfortunately indicated a high level of ribosomal contamination. To further reduce nonspecific interactions in the purified particle, a third stage was added to the purification. This step utilized the calmodulin binding domain of the TAP tagged Prp6p. In brief, the peak glycerol gradient fractions of 15-19 were pooled and applied to a column filled with calmodulin resin. A five fraction EGTA elution was followed by RNA extraction and ultimately northern blot analysis. This revealed that that all five snRNAs remained intact in the particle (Figure 4.10).

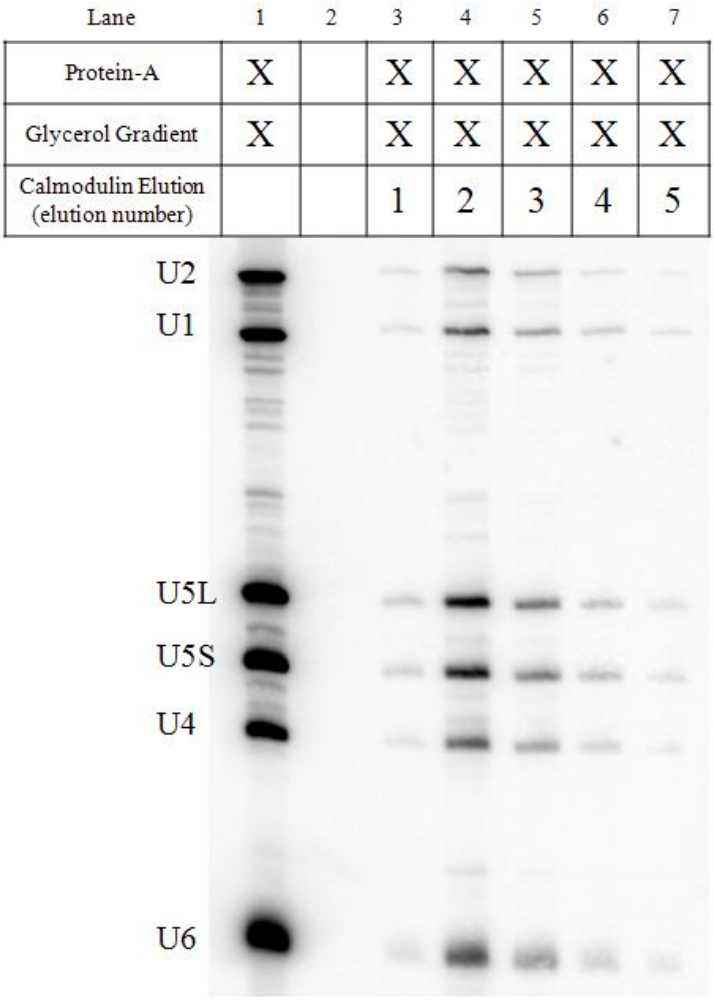


Figure 4.10 - All five snRNAs are retained during a three-step purification

Northern blot analysis of RNA fraction eluate during a two and three step purification. After the first IgG affinity step, the contents were placed on a glycerol gradient. The peak fractions from the gradient were pooled and a sample was saved for later analysis (Lane 1). The remainder of the material was loaded, washed, and eluted from a calmodulin column in five fractions (Lane 3-7).

4.2.7 Mass Spectrometric Analysis of RNP Complex

In order to obtain adequate quantities of the particle for mass spectrometry, five IgG affinity purifications were performed in parallel. The eluate from each of these columns was sedimented through a glycerol gradient and fractionated. Based on the results shown in Figure 4.2, fractions 15 - 19 were pooled from each of the gradients. This material was loaded to five calmodulin columns in parallel. Upon elution from the column with EGTA, the material was phenol chloroform extracted and the organic layer acetone precipitated as in section 4.2.6. After pelleting any proteins present, the sample was submitted to mass spectrometric analysis undertaken with collaborators at the Yates lab at the Scripps Research Institute, CA. This analysis utilized a system termed Automated Ultra-High Pressure Multidimensional Protein Identification Technology (UHP-MudPIT) (Motoyama, Venable, Ruse, and John R Yates 2006) that relies on liquid chromatography to separate the single sample complex protein mixture prior to analysis.

Examination of the data revealed that proteins from all five snRNPs were present, along with proteins from the Prp19 complex and numerous splicing-associated factors (Table 4.1).

Protein	Systematic Name	Molecular Mass (kDa)
U1 Proteins		
PRP42	YDR235W	65
SNU71	YGR013W	71
PRP40	YKL012W	69
PRP39	YML046W	75
NAM8	YHR086W	57
SNU56	YDR240C	56
NPL3	YDR432W	45
SNP1	YIL061C	34
YHC1	YLR298C	27
LUC7	YDL087C	30
MUD1	YBR119W	34
U2 Proteins		
RSE1	YML049C	154
HSH155	YMR288W	110
CUS1	YMR240C	50
PRP9	YDL030W	63
PRP21	YJL203W	33
PRP11	YDL043C	30
HSH49	YOR319W	25
IST3	YIR005W	17
LEA1	YPL213W	27
MSL1	YIR009W	13
YSF3	YNL138W-A	10
U4/U6.U5		
PRP8	YHR165C	280
BRR2	YER172C	246
SNU114	YKL173W	114
PRP6	YBR055C	104
PRP31	YGR091W	56
PRP3	YDR473C	56
PRP4	YPR178W	52
SPP381	YBR152W	34
PRP28	YDR243C	67
SNU23	YDL098C	23
DIB1	YPR082C	17
SNU13	YEL026W	14

Protein	Systematic Name	Molecular Mass (kDa)
NTC		
PRP19	YLL036C	57
CEF1	YMR213W	68
PRP46	YPL151C	51
SYF1	YDR416W	100
SNT309	YPR101W	21
ISY1	YJR050W	28
SYF2	YGR129W	25
CWC2	YDL209C	38
Sm/Lsm		
SmB1	YER029C	22
SmD1	YGR074W	16
SmD2	YLR275W	13
SmD3	YLR147C	11
SmF	YPR182W	10
SmG	YFL017W-A	8
LSM1	YJL124C	20
LSM2	YBL026W	11
LSM4	YER112W	21
LSM5	YER146W	10
LSM6	YDR378C	9
LSM7	YNL147W	13
LSM8	YJR022W	12
LSM12	YHR121W	21
mRNA Binding		
PAB1	YER165W	64
NPL3	YDR432W	45
GLC7	YER133W	36
YRA1	YDR381W	25
STO1	YMR125W	100
CBC2	YPL178W	24
YRA2	YKL214C	24
ABD1	YBR236C	50
BUD13	YGL174W	30
GLE1	YDL207W	62
REF2	YDR195W	60
NRD1	YNL251C	64
MUD2	YKL074C	60
JSN1	YJR091C	120
PBP1	YGR178C	79

Protein	Systematic Name	Molecular Mass (kDa)
Other Splicing		
DED1	YOR204W	66
DHH1	YDL160C	58
SPP2	YOR148C	21
PRP2	YNR011C	100
YJU2	YKL095W	32
NTR1	YLR424W	83
NTR2	YKR022C	37
SLU7	YDR088C	45
EXO84	YBR102C	86
SUB2	YDL084W	50
KEM1	YGL173C	175
PRP43	YGL120C	88
Possible Contaminants		
AHP1	YLR109W	19
PGK1	YCR012W	45
ENO1	YGR254W	47
TSA1	YML028W	22
ADH1	YOL086C	37
FPR1	YNL135C	12
TEF1	YPR080W	50
TEF2	YBR118W	50
HSC82	YMR186W	81
SSE1	YPL106C	77
SBP1	YHL034C	33
ACT1	YFL039C	42
IPP1	YBR011C	32
UBP3	YER151C	102
FAS2	YPL231W	207
YDR161W	YDR161W	43
SKI2	YLR398C	146

Table 4.1 - Results of mass spectrometry on the A2-1 complex

Mass Spectrometry analysis of an affinity-purified 55S particle reveal proteins from all five snRNPs as well as the PRP19 complex. After IgG affinity purification followed by glycerol gradient separation, peak fractions were further purified via calmodulin affinity chromatography. Proteins from this column's eluate were precipitated and analyzed by MudPIT mass spectrometry.

4.3 DISCUSSION

As described in Chapter 3, cold-shifted *Δsnu40+Δsnu66* double deletion strains exhibit a buildup of a spliceosome-sized particle containing Prp6p. The results revealed that Snu40p and Snu66p were responsible for the handoff of Prp6p from the U5 snRNP to the newly forming U4/U6·U5 tri-snRNP. When *SNU40* and *SNU66* were deleted from the cell, the position of Prp6p in the tri-snRNP became destabilized. Additionally, Brr2p was greatly reduced in the ability to unwind U4 from U6.

From these results, the identity of the large Prp6p containing complex could be inferred. Our hypothesis was that after tri-snRNP formation, the spliceosome was assembling typically. However, at reduced temperatures, the inability of Brr2p to unwind U4 from U6 would result in a stalled precatalytic (A2-1) complex. With this hypothesis in mind, we set about the purification of the particle in question.

Initially we started the purification as a two step process. The first step involved a simple affinity purification via an IgG Sepharose interaction with a TAP-tagged Prp6p. This was followed by elution from the column by TEV protease and glycerol gradient separation of the resulting complexes. It is worth noting that in an attempt to retain condition similarity with the extract in which the particle was originally identified, the purification was performed under conditions typical (Ansari and Schwer 1995) to those used in splicing reactions. In short, care was taken to harvest the cells at 4°C with extensive washes. Cell disruption utilized a cryogenic ball mill, which disrupts cells in a manner similar to a mortar and pestle. After cell thawing and a high salt extraction phase, all lysate was adjusted to splicing-compatible 50 mM NaCl before proceeding with the purification.

The resulting material, when extracted and precipitated, revealed all five snRNAs to be present in a complex approximately 55S in size (Figure 4.2, fractions 15-19). Additionally, a significant amount of particle containing the U4, U5, and U6 snRNAs was seen in fractions 7-9. While this species may be the product of a native process, the relative lack of accumulated tri-snRNP seen in Figure 4.1 raises the alternate possibility that this is an experimental artifact resulting from destabilization of the larger particle found lower in the gradient.

In addition to the snRNA composition, the base-pairing status of U2, U4, and U6 aids in identifying the particle in question. Prior to spliceosome activation, U4 and U6 remain extensively basepaired. Upon activation and progression from the A2-1 complex to the A1 complex U4 is unwound from U6 which then addresses the 5' splice site. U6 snRNA concurrently forms a 11nt basepairing arrangement with the U2 snRNA (D.A. Wassarman and J.A. Steitz 1993). With these options in mind, we identified the status of these partners in the purified 55S complex. The first method used to determine basepairing endeavored to identify the base-pairing status of the U6 snRNA. Prior experience in the Stevens lab (unpublished data) indicates that phenol-chloroform extraction and ethanol precipitation of RNA is not does not perturb a paired U4/U6 duplex. Separating the RNA from the 55S gradient fractions under native gel conditions followed by northern blotting for U6 revealed that 100% of the snRNA is in a basepaired state. This experiment was then repeated with the addition of a psoralen crosslinking technique. After crosslinking the particle, the RNA was extracted and precipitated. The sample was then split and separated under denaturing conditions on a urea-acrylamide gel. Although the crosslinked RNA was run in three separate lanes, they were all run on the same gel thus assuring comparable migration rates. This gel was then dissected and transferred for northern blotting. The northern blot results revealed that the U2 snRNA

barely entered the gel, most likely due to a crosslink to the pre-mRNA and its large size. The co-migration of U4 and U6 indicate that they are crosslinked to each other and moreover are not crosslinked to the mRNA. Thus, the basepairing status of U4 and U6 indicates the presence of an A2-1 complex.

While previous work (Scott W. Stevens et al. 2002) in this lab indicated that all five snRNAs can be purified in the absence of a substrate pre-mRNA, the A2-1 complex exists in the presence of a substrate. To distinguish between these two possibilities we selected a few highly transcribed RNAs for analysis. RT-PCR of the *RPS4A*, *RPS10A*, and *RPL10A* transcripts indicate the presence of high amounts of pre-mRNA. With this, we are able to conclude with certainty that we have purified an A2-1 splicing complex.

Although we have successfully identified this purified particle, it is formally possible that it was a nonfunctional dead-end product resulting from the deletion of *SNU40* and *SNU66*. If the complex were simply stalled due to the U4/U6 unwinding defect at low temperatures there should be minimal effect on splicing once the particle is activated. Alternately, the deletions may have downstream effects on splicing that are independent of U4 and U6 unwinding. To distinguish between these hypotheses, we tested the purified particle by two different assays. The first approach was to assess the ability via RT-PCR of the A2-1 complex to complete splicing of its bound substrate. The second technique was intended to assess the ability of this complex and its components to recycle through multiple rounds of splicing. The second method utilized the traditional radiolabeled *in vitro* transcribed pre-mRNA splicing assay.

There are two theoretical advantages of the first listed approach in analyzing the splicing efficiency of the substrate pre-mRNA. The first advantage of this system results from the use of an *in vivo* assembled substrate. The second advantage is the ability to analyze differences and similarities in splicing efficiency between any number of

transcript species while in a defined *in vitro* setting. While our current sample size of three is obviously inadequate for large scale differential transcript analysis, future work will leverage this platform. Additionally it should be noted that the RT-PCR data exhibited in this dissertation is preliminary in nature. As such, the results seen here are semi-quantitative and will be followed with multiple replicates of real time RT-PCR analysis prior to publication. As is, we are reasonably able to conclude that our purified A2-1 complex is indeed capable of completing the splicing cycle when in the presence of exogenous protein factors. Lastly, it is noteworthy that an apparent increase in lariat intermediate species indicates that the particle may not function in this assay with upmost efficiency.

Although the *in vivo* assembled particle appears to be deficient in progressing through the second step of splicing, it is a formal possibility that this defect is the result of either snRNA or protein component defects. To distinguish between these options we carried out a splicing reaction of exogenous in vitro pre-mRNA in a RNA free micrococcal nuclease treated extract. As spliceosome components are recycled they should readdress the freshly added pre-mRNA. Once recycled, however, the snRNAs should be complimented with the complete set of splicing proteins.

As increasing amounts of the particle were added to the reaction we observed that correspondingly increasing amounts of mature mRNA product was being produced. Quantitation analysis revealed that when additional volumes of particle were placed into the reaction, an increase in completed mRNA product as well as intermediate species could be observed. Thus, the snRNAs that were purified were indeed able to recycle through multiple rounds of splicing.

Given the positive identification of this stalled A2-1 complex and subsequent assessment of its catalytic viability, the obvious remaining challenge was to identify the

components of this complex. We were fortunate to have willing collaborators in the Yates lab at the Scripps Research Institute. Their expertise and highly sensitive equipment made analysis of the protein components by mass spectrometry an efficient process.

The coomassie stained gel of the glycerol gradient fractions (Figure 4.3) revealed numerous proteins present in the fractions corresponding with those containing snRNAs. When the protein fractions of the 55S complex were examined by mass spectrometry, however, they revealed a high level of background contamination. To reduce these contaminants, future gradient material was passed over a second affinity column, this time utilizing the calmodulin binding domain of the TAP tag. Combined with a final high salt (250 mM) wash while on the calmodulin column, the results were far superior to earlier attempts.

The mass spectrometry analysis revealed that components from all five snRNPs were present. 94% (16/17) of established U1 proteins were found, along with 85% (17/20) of known U2 proteins. Of the core tri-snRNP proteins, 82% (23/28) were found in the particle. Additionally, all eight members of the Prp19p associated complex (NineTeen Complex, NTC) were identified. Most of the members of the Sm and Lsm class were identified, with the exception of Lsm3 and SmE. Other proteins of interest that were identified in the complex were Spp2, Prp2, and Yju2.

The characterization of a stalled A2-1 complex sheds significant light on the assembly of the Prp19p related complex (Chun-Hong Chen et al. 2002) and associated members. Not only are the functions of this complex still being elucidated, but it remains unclear at what stage the components join the spliceosome. Previous work (Chan, Der-I Kao, Wei-Yu Tsai, and Soo-Chen Cheng 2003) (Chan and Soo-Chen Cheng 2005) indicates that the NTC is required for stable association of U5 and U6 in the post-U4 spliceosome. Additional work suggested that the NTC joined the spliceosome near the

time of U4 disassociation while remaining noncommittal as to the specific order of these events (Chun-Hong Chen et al. 2002). The presence of the complete NTC in a stalled A2-1 complex in this dataset, solves this quandary.

The DEAH box helicase Prp2p is responsible for the rearrangements required for the first transesterification reaction of splicing (S.H. Kim and R.J. Lin 1993). Spp2p is established as interacting with the spliceosome prior to the joining of Prp2p (Roy, K. Kim, Maddock, Anthony, and Jr 1995). Further, Spp2p interaction with the C-terminus of Prp2 is required for the helicase's activity (Silverman et al. 2004). In spite of what is known about these proteins, it remains unknown when exactly they join the spliceosome. Based on the results here, it can be concluded that Spp2p and Prp2p join the spliceosome prior to activation. Of the three other proteins shown to interact with Prp2p (Cef1 (Gavin et al. 2002), Cin2 (Ito, Chiba, Ozawa, Yoshida, Masahira Hattori, and Yoshiyuki Sakaki 2001), Brr2p (van Nues and Jean D. Beggs 2001)) two of them (Cef1, Brr2) were isolated in the A2-1 complex.

The NTC associated Yju2p has been shown to function after the Prp2-dependent spliceosomal rearrangement to promote the first step of splicing (Yen-Chi Liu, Hsin-Chou Chen, Nan-Ying Wu, and Soo-Chen Cheng 2007). While it is known when Yju2p acts in the splicing cycle, it is unknown when it joins the complex. Competing theories exist that place Yju2p in the spliceosome with the NTC or alternately after activation (A1 complex formation). Our data demonstrate that Yju2p has indeed joined the spliceosome prior to A2-1 complex formation.

Although fewer data points on spliceosome function have been gathered in the mammalian system, it is notable that the counterpart of the A2-1 complex has recently been characterized (Deckert, Hartmuth, Boehringer, Behzadnia, Cindy L. Will, Berthold Kastner, Holger Stark, Henning Urlaub, and Reinhard Luhrmann 2006). Known as the B

complex, the precatalytic mammalian spliceosome is also the only snRNP complex known to contain all five snRNAs in the presence of a substrate pre-mRNA. Complicating the comparison of these two datasets, however, are two issues. The first issue is that of missing homologs. As there are significant evolutionary differences between these species, differences in protein content and function are to be expected in spite of the generally high level of conservation in the core spliceosomal machinery. The second complicating factor in comparing data sets is that of naming convention. For a variety of reasons homologous proteins differ in name between these two species. To aid in the comparison of the *Saccharomyces cerevisiae* and human datasets, a translation table of proteins known to be of importance in splicing. It is included in Appendix I.

Overall, the two datasets show a high degree of alignment (Appendix II). There are a number of proteins found in the *Saccharomyces cerevisiae* spliceosome that have no known homolog in the mammalian system. Ignoring these differences, there are only three proteins found in our core spliceosome dataset (Table 1.1) that was are not also confirmed by the mammalian data (Snu17p, Prp4p, Lsm3) (Detail in Appendix II). The comparison of identified mRNA binding proteins as well as splicing associated factors has a less complete overlap, however, with 24 proteins remaining unconfirmed by the between the two datasets.

As pointed out elsewhere (Maria M Konarska 2008), what was considered to be the B complex in this work (Deckert et al. 2006) is actually the combination of at least three different species. Criticism of this work has noted that this heterogeneity may account for the appearance of the mammalian NTC counterpart - as it was generally thought to join the spliceosome after the exit of U1 and U4. On the basis of U4/U6 binding status and stoichiometric snRNA ratios, our data would not appear to suffer from this same heterogeneity. The A2-1 complex purified here is a mono species including all

five snRNAs that results from the *Δsnu40* + Δ*snu66* induced U4/U6 unwinding defect. The presence of contaminating mature mRNA would seem to complicate these findings, however. Given that the vast majority of cellular mRNA has already been spliced and the exquisite sensitivity of the RT-PCR assay, the presence of detectable levels of this contaminant is not surprising. It should be pointed out that the presence of this spliced product does not impact the RT-PCR data of the pre-mRNA nor of the intron-exon 2 product. Further, in analyzing the splicing efficiency of the purified material, one can assume that the substrate RNA is conserved through the splicing process. Any analysis that tracks each species as a percentage of the total at any point in time will capture net migration regardless of the starting volumes. In any event, our data on the NTC remains particularly compelling when the mammalian B complex data is considered. Indeed, in combination with this information we are able to establish that the NTC and related factors are present in the spliceosome prior to the unwinding of U4 and U6.

In addition to the comparison of the A2-1 complex and the earlier mammalian B complex, Appendix II also includes a comparison with earlier data (Scott W. Stevens et al. 2002) on the penta-snRNP complex. It is the overall similarity that stands out in comparing the core components of the A2-1, B, and penta-snRNP complexes. Almost all of the core U1, U2, tri-snRNP and NTC proteins are present in the A2-1 / B complex and penta-snRNP. Some of the remaining differences are likely the result of peptide coverage issues. For example, the missing SmE in the A2-1 and B complexes is unlikely to be a true biological result. On the other hand, it is conceivable that some of the other differences represent actual differences between homologous species. It is difficult to predict which of these are worthy of further scrutiny, but the simple illumination of the variation is surely the first step of analysis.

In moving beyond the core snRNP proteins, true differences between the penta-snRNP and A2-1/B complex stand out. Fifteen proteins identified as being associated with mRNA binding were identified in the A2-1 complex. Of these, almost half were identified in the B complex. In contrast, none were identified in the penta-snRNP. Of course this result is to be expected as the penta-snRNP lacks the substrate mRNA that carries these proteins.

As discussed earlier, the NTC associated protein Yju2p was identified in the A2-1 complex. This protein was not found in the penta-snRNP complex, nor was the mammalian counterpart (CCDC130) found in the B complex. A similar situation exists for Spp2p which was identified in A2-1 to the exclusion of the other two complexes. The A2-1 complex data stands in particular opposition to the B complex in this last case. Given that the B complex successfully identified the mammalian Prp2p homolog, the lack of the mammalian Spp2 (GPKOW) seems unexpected. As Spp2p is thought to join the spliceosome prior to its interaction with Prp2p, one would not expect Prp2p to be found exclusively. In contrast, both proteins were found in the A2-1 complex.

APPENDIX I – SPLICING PROTEIN HOMOLOGS

Saccharomyces Cerevisiae		Homo sapiens	
Protein Name	**Association**	**Protein Name**	**Association**
Luc7p	U1	-	-
Mud1p	U1	A	U1
Nam8p	U1	-	-
Prp39p	U1	-	-
Prp40p	U1	-	-
Prp42p	U1	-	-
Snp1p	U1	70K	U1
Snu56p	U1	-	-
Snu71p	U1	-	-
Yhc1p	U1	U1C	U1
Cus1p	U2	SF3b145	U2
Cus2p	U2	-	-
Hsh155p	U2	SF3b155	U2
Hsh49p	U2	SF3b49	U2
Lea1p	U2	A'	U2
Msl1p	U2	B''	U2
Prp11p	U2	SF3a66	U2
Prp21p	U2	SF3a120	U2
Prp9p	U2	SF3a60	U2
Rse1p	U2	SF3b130	U2
Snu17p	U2	CGI-79	RES complex
Prp43p	U2/disassembly	hPrp43	U2
Prp5p	B complex	hPrp5	U2
Ysf3p	U2	SF3b10	U2
Rds3p	U2 associated	SF3b14b	U2
-	-	SF3b125	U2
-	-	SF3b14a	U2
-	-	SR140	U2
-	-	CHERP	U2
-	-	SPF45	U2
-	-	SPF31	U2
-	-	SPF30	U2

Saccharomyces Cerevisiae		Homo sapiens	
Protein Name	**Association**	**Protein Name**	**Association**
Prp3p	U4/U6	90K	U4/U6
Prp31p	U4/U6	61K	U4/U6
Prp4p	U4/U6	60K / hPRP4	U4/U6
Snu13p	U4/U6	15.5K	U4/U6
-	-	20K	U4/U6
Prp24p	U6	SART3 / p110	U4/U6
Prp8p	U5	220K	U5
Brr2p	U5	200K	U5
Snu114p	U5	116K	U5
Prp28p	U5	100K	U5
Snu40p	U5	52K	U5
Dib1p	U5	15K	U5
Aar2p	U5	-	-
Prp6p	U5	102K	U5
Prp3p	U5	90K	U4/U6
Prp18p	U5	-	-
-	-	40K	U5
Prp38p	U4/U6•U5	hPrp38	U4/U6•U5
Sad1p	U4/U6•U5	65K	U4/U6•U5
Snu23p	U4/U6•U5	-	-
Snu66p	U4/U6•U5	110K / SART1	U4/U6•U5
Spp381p	U4/U6•U5	-	-
-	-	27K	U4/U6•U5
Lsm2p	U6	Lsm2	U6
Lsm3p	U6	Lsm3	U6
Lsm4p	U6	Lsm4	U6
Lsm5p	U6	Lsm5	U6
Lsm6p	U6	Lsm6	U6
Lsm7p	U6	Lsm7	U6
Lsm8p	U6	Lsm8	U6

Saccharomyces Cerevisiae		Homo sapiens	
Protein Name	**Association**	**Protein Name**	**Association**
Smb1p	U1, U2, U4, U5	Smb1	U1, U2, U4, U5
Smd1p	U1, U2, U4, U5	Smd1	U1, U2, U4, U5
Smd2p	U1, U2, U4, U5	Smd2	U1, U2, U4, U5
Smd3p	U1, U2, U4, U5	Smd3	U1, U2, U4, U5
Sme1p	U1, U2, U4, U5	Sme1	U1, U2, U4, U5
SmFp	U1, U2, U4, U5	SmF	U1, U2, U4, U5
SmGp	U1, U2, U4, U5	SmG	U1, U2, U4, U5
Prp19p	NTC	hPRP19	Prp19 Complex
CEF1p	NTC	Cdc5	Prp19 Complex
PRP46p	NTC	PRL1	Prp19 Complex
SYF1p	NTC	hSYF1 (XAB2)	Prp19 Complex
SNT309p	NTC	SPF27	Prp19 Complex
ISY1p	NTC	hIsy1 (fSAP133)	Prp19 Complex
SYF2p	NTC	SYF2	Prp19 Complex
CWC2p	NTC	-	Prp19 Complex
Slt11p	Spliceosome	ECM2	Spliceosome
Slu7p	Spliceosome / 2nd Step	-	-
Pab2p	Cap binding	PAB2	Cap binding
Mud2p	Spliceosome / 1st Step	U2AF65	Spliceosome / 1st Step
Prp16p	Spliceosome / 2nd Step	PRP16	Spliceosome / 2nd Step
Prp17p	Spliceosome / 2nd Step	PRP17	Spliceosome / 2nd Step
Prp22p	Spliceosome / 2nd Step	PRP22	Spliceosome / 2nd Step
Prp45p	Spliceosome / 2nd Step	SKIP	Spliceosome / 2nd Step
Dbp2p	Unknown	p68	U1 Associated
Prp2p	Spliceosome / 1st Step	hPRP2	Spliceosome / 1st Step
Bbp1p	Branchpoint	mBBP/SF1	Branchpoint
Sub2p	Prespliceosome	hUAP56	Prespliceosome
Spp2p	U2 associated	GPKOW	U2 associated
Yju2p	NTC associated	CCDC130	NTC associated
Xrn1	penta-snRNP	Xrn1	exonuclease

Data based on: (Berglund, Abovich, and Rosbash 1998; Libri, Graziani, Saguez, and Boulay 2001; Bell, Schreiner, Damianov, Ram Reddy, and Bindereif 2002; Zhi-Ren Liu 2002; Jurica and M.J. Moore 2002; Vincent, Q. Wang, Jay, Hobbs, and B.C. Rymond 2003; Qiang Wang and Brian C Rymond 2003; Sunbin Liu 2005; Lesley Collins and Penny 2005; Benson, Karsch-Mizrachi, Lipman, Ostell, and David L Wheeler 2006; Deckert et al. 2006)

APPENDIX II – PROTEINS OF THE A2-1, PENTA-SNRNP, AND MAMMALIAN B COMPLEX

Protein	Yeast A2-1 Complex	Yeast penta-snRNP	Mammalian B Complex
U1 Proteins			
PRP42p	X	X	N/A
SNU71p	X	X	N/A
PRP40p	X	X	X
PRP39p	X	X	N/A
NAM8p	X	X	N/A
SNU56p	X	X	N/A
NPL3p	X		
SNP1p	X	X	X
YHC1p	X	X	X
LUC7p	X	X	N/A
MUD1p	X	X	X
U2 Proteins			
RSE1p	X	X	X
HSH155p	X	X	X
CUS1p	X	X	X
PRP9p	X	X	X
PRP21p	X	X	X
PRP11p	X	X	X
HSH49p	X	X	X
IST3p	X		X
LEA1p	X	X	X
MSL1p	X	X	X
YSF3p	X		X
Cus2p		X	
Snu17p		X	X

Protein	Yeast A2-1 Complex	Yeast penta-snRNP	Mammalian B Complex
U4/U6.U5			
PRP8p	X	X	X
BRR2p	X	X	X
SNU114p	X	X	X
PRP6p	X	X	X
PRP31p	X	X	X
PRP3p	X	X	X
PRP4p	X		X
SPP381p	X	X	N/A
PRP28p	X		X
SNU23p	X	X	X
DIB1p	X	X	X
SNU13p	X	X	X
Snu66	N/A	X	X
Prp4b		X	X
Prp38		X	X
NTC			
PRP19p	X	X	X
CEF1p	X	X	X
PRP46p	X	X	X
SYF1p	X	X	X
SNT309p	X	X	X
ISY1p	X	X	X
SYF2p	X	X	X
CWC2p	X		N/A
Syf3		X	X

Protein	Yeast A2-1 Complex	Yeast penta-snRNP	Mammalian B Complex
Sm/Lsm			
SmB1p	X	X	X
SmD1p	X	X	X
SmD2p	X	X	X
SmD3p	X	X	X
SmEp		X	
SmFp	X	X	X
SmGp	X	X	X
LSM1p	X	X	
LSM2p	X	X	X
Lsm3p		X	X
LSM4p	X	X	X
LSM5p	X		X
LSM6p	X	X	X
LSM7p	X	X	X
LSM8p	X	X	X
LSM12p	X		
mRNA Binding			
PAB1p	X		X
NPL3p	X		
GLC7p	X		X
YRA1p	X		X
STO1p	X		X
CBC2p	X		X
YRA2p	X		
ABD1p	X		
BUD13p	X		X
GLE1p	X		
REF2p	X		
NRD1p	X		
MUD2p	X		X
JSN1p	X		
PBP1p	X		

Protein	Yeast A2-1 Complex	Yeast penta-snRNP	Mammalian B Complex
Other Splicing			
DED1p	X	X	X
DHH1p	X	X	
SPP2p	X		
PRP2p	X		X
YJU2p	X		
NTR1p	X		X
NTR2p	X		
SLU7p	X		X
EXO84p	X		
SUB2p	X		X
KEM1p	X		
PRP43p	X		X
MSL5p	X		
Swi4		X	
Mbp1		X	
Swi6		X	
Xm1		X	
Pat1		X	
Ssa		X	X
Ssb		X	
Asc1		X	
Penta-snRNP			
Sad1		X	X
Prp45		X	X
Ecm2		X	X
Cwf2		X	
Ntc20		X	

Protein	Yeast A2-1 Complex	Yeast penta-snRNP	Mammalian B Complex
Other			
Pur5		X	
YLR432W		X	
YML056C		X	
YAR073W		X	
Tef1	X	X	
Nop1		X	
Sis1		X	
Tdh1		X	
Gpm1		X	
YLR016C		X	
YCR063W		X	
YDR428C		X	
Rps0e		X	
Rsp1		X	
Rsp2		X	
Rsp3		X	
Rsp4		X	
Rps5		X	
Rsp7		X	
Rsp9		X	
Rsp12		X	
Rsp13		X	
Rsp14		X	
Rps16		X	
Rps18		X	
Rsp19		X	
Rsp20		X	
AHP1	X		
PGK1	X		
ENO1	X		
TSA1	X		
ADH1	X		
FPR1	X		
TEF2	X		
HSC82	X		
SSE1	X		
SBP1	X		
ACT1	X		
IPP1	X		
UBP3	X		
FAS2	X		
YDR161W	X		
SKI2	X		

REFERENCES

Abovich, N., Legrain, P., and Rosbash, M. 1990. The yeast PRP6 gene encodes a U4/U6 small nuclear ribonucleoprotein particle (snRNP) protein, and the PRP9 gene encodes a protein required for U2 snRNP binding. *Molecular and cellular biology* **10**: 6417-25.

Achsel, T., Ahrens, K., Brahms, H., Teigelkamp, S., and Lührmann, R. 1998. The human U5-220kD protein (hPrp8) forms a stable RNA-free complex with several U5-specific proteins, including an RNA unwindase, a homologue of ribosomal elongation factor EF-2, and a novel WD-40 protein. *Mol. Cell. Biol.* **18**: 6756-6766.

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., and Walter, P. 2002. *Molecular Biology of the Cell, Fourth Edition*. 4th ed. Garland.

Ansari, A., and Schwer, B. 1995. SLU7 and a novel activity, SSF1, act during the PRP16-dependent step of yeast pre-mRNA splicing. *EMBO J.* **14**: 4001-9.

Azubel, M., Habib, N., Sperling, R., and Sperling, J. 2006. Native spliceosomes assemble with Pre-mRNA to form supraspliceosomes. *J. Mol. Biol.* **356**: 955-966.

Azubel, M., Wolf, S., Sperling, J., and Sperling, R. 2004. Three-dimensional structure of the native spliceosome by cryo-electron microscopy. *Molecular Cell* **15**: 833-839.

Bach, M., Winkelmann, G., and Lührmann, R. 1989. 20S small nuclear ribonucleoprotein U5 shows a surprisingly complex protein composition. *Proc. Natl. Acad. Sci.* **86**: 6038-6042.

Baron, S. 1996. *Medical Microbiology*. 4th ed. University of Texas Medical Branch.

Behrens, S., and Lührmann, R. 1991. Immunoaffinity purification of a [U4/U6•U5] tri-snRNP from human-cells. *Genes Dev.* **5**: 1439-1452.

Behzadnia, N., Hartmuth, K., Will, C.L., and Luhrmann, R. 2006. Functional spliceosomal A complexes can be assembled in vitro in the absence of a penta-snRNP. *RNA* **12**.

Bell, M., Schreiner, S., Damianov, A., Reddy, R., and Bindereif, A. 2002. p110, a novel human U6 snRNP protein and U4/U6 snRNP recycling factor. *The EMBO journal* **21**: 2724-35.

Benson, D.A., Karsch-Mizrachi, I., Lipman, D.J., Ostell, J., and Wheeler, D.L. 2006. GenBank. *Nucleic acids research* **34**: D16-20.

Berg, J.M., Tymoczko, J.L., and Stryer, L. 2006. *Biochemistry*. 6th ed. W. H. Freeman.

Berget, S.M., Moore, C., and Sharp, P.A. 1977. Spliced segments at the 5' terminus of adenovirus 2 late mRNA. *Proceedings of the National Academy of Sciences of the United States of America* **74**: 3171-5.

Berglund, J., Abovich, N., and Rosbash, M. 1998. A cooperative interaction between U2AF65 and mBBP/SF1 facilitates branchpoint region recognition. *Genes Dev.* **12**: 858-867.

Biemann, K., and Scoble, H.A. 1987. Characterization by tandem mass spectrometry of structural modifications in proteins. *Science (New York, N.Y.)* **237**: 992-8.

Birnstiel, M.L. 1988. *Structure and Function of Major and Minor Small Nuclear Ribonucleoprotein Particles*. Springer.

Boeke, J., Trueheart, J., Natsoulis, G., and Fink, G. 1987. 5-Fluoroorotic Acid As a Selective Agent in Yeast Molecular- Genetics. *Method Enzymol.* **154**: 164-175.

Borman, S., Russell, H., and Siuzdak, G. 2003. A Mass Spectrometry Timeline. *Today's Chemist* 47-49.

Botstein, D., and Fink, G. 1988. Yeast: an experimental organism for modern biology. *Science* **240**: 1439-1443.

Brakke, M.K. 1953. Zonal separations by density-gradient centrifugation. *Archives of biochemistry and biophysics* **45**: 275-90.

Branlant, C., Krol, A., Ebel, J.P., Lazar, E., Gallinaro, H., Jacob, M., Sri-Widada, J., and Jeanteur, P. 1980. Nucleotide sequences of nuclear U1A RNAs from chicken, rat and man. *Nucleic acids research* **8**: 4143-54.

Busch, H., Reddy, R., Rothblum, L., and Choi, Y.C. 1982. SnRNAs, SnRNPs, and RNA processing. *Annual review of biochemistry* **51**: 617-54.

Carr, S.A., Hemling, M.E., Bean, M.F., and Roberts, G.D. 1991. Integration of mass spectrometry in analytical biotechnology. *Analytical Chemistry* **63**: 2802-2824.

Chan, S., and Cheng, S. 2005. The Prp19-associated Complex Is Required for Specifying Interactions of U5 and U6 with Pre-mRNA during Spliceosome Activation. *J. Biol. Chem.* **280**: 31190-31199.

Chan, S., Kao, D., Tsai, W., and Cheng, S. 2003. The Prp19p-Associated Complex in Spliceosome Activation. *Science* **302**: 279-282.

Chanfreau, G., Elela, S.A., Ares, M., and Guthrie, C. 1997. Alternative 3'-end processing of U5 snRNA by RNase III. *Genes & development* **11**: 2741-51.

Chapman, K., and Boeke, J. 1991. Isolation and characterization of the gene encoding yeast debranching enzyme. *Cell* **65**: 483-92.

Chen, C., Tsai, W., Chen, H., Wang, C., and Cheng, S. 2001. Identification and characterization of two novel components of the prp19p-associated complex, Ntc30p and Ntc20p. *Journal Of Biological Chemistry* **276**: 488-494.

Chen, C., Yu, W., Tsao, T.Y., Wang, L., Chen, H., Lin, J., Tsai, W., and Cheng, S. 2002. Functional and physical interactions between components of the Prp19p-associated complex. *Nucl. Acids Res.* **30**: 1029-1037.

Chen, E.I., Hewel, J., Felding-Habermann, B., and Yates, J.R. 2006. Large Scale Protein Profiling by Combination of Protein Fractionation and Multidimensional Protein Identification Technology (MudPIT). *Mol Cell Proteomics* **5**: 53-56.

Chen, H., Jan, S., Tsao, T.Y., Sheu, Y., Banroques, J., and Cheng, S. 1998. Snt309p, a Component of the Prp19p-Associated Complex That Interacts with Prp19p and Associates with the Spliceosome Simultaneously with or Immediately after Dissociation of U4 in the Same Manner as Prp19p. *Mol. Cell. Biol.* **18**: 2196-2204.

Chen, J., and Lin, R. 1990. The yeast PRP2 protein, a putative RNA-dependent ATPase, shares extensive sequence homology with two other pre-mRNA splicing factors. *Nucleic Acids Res.* **18**: 6447.

Chen, Y.G., Moore, R.E., Ge, H.Y., Young, M.K., Lee, T.D., and Stevens, S.W. 2007. Proteomic analysis of in vivo-assembled pre-mRNA splicing complexes expands the catalog of participating factors. *Nucleic acids research* **35**: 3928-44.

Cheng, S., and Abelson, J. 1987. Spliceosome assembly in yeast. *Genes Dev.* **1**: 1014-1027.

Cheng, S., Tarn, W., Tsao, T., and Abelson, J. 1993. Prp19 - a Novel Spliceosomal Component. *Mol. Cell. Biol.* **13**: 1876-1882.

Cho, E., Takagi, T., Moore, C.R., and Buratowski, S. 1997. mRNA capping enzyme is recruited to the transcription complex by phosphorylation of the RNA polymerase II carboxy-terminal domain. *Genes Dev.* **11**: 3319-3326.

Chow, L.T., Gelinas, R.E., Broker, T.R., and Roberts, R.J. 1977. An amazing sequence arrangement at the 5' ends of adenovirus 2 messenger RNA. *Cell* **12**: 1-8.

Church, G., and Gilbert, W. 1984. Genomic Sequencing. *Proc. Natl. Acad. Sci. U.S.A.* **81**: 1991-1995.

Collins, L., and Penny, D. 2005. Complex Spliceosomal Organization Ancestral to Extant Eukaryotes. *Mol Biol Evol* **22**: 1053-1066.

de la Cruz, J., Kressler, D., and Linder, P. 1999. Unwinding RNA in Saccharomyces cerevisiae: DEAD-box proteins and related families. *Trends Biochem.Sci.* **24**: 192-198.

Darnell, J., Matsudaira, P., Zipursky, L., Lodish, H., Berk, A., and Baltimore, D. 1999. *Molecular Cell Biology*. 4th ed. W. H. Freeman.

Deckert, J., Hartmuth, K., Boehringer, D., Behzadnia, N., Will, C.L., Kastner, B., Stark, H., Urlaub, H., and Luhrmann, R. 2006. Protein Composition and Electron Microscopy Structure of Affinity-Purified Human Spliceosomal B Complexes Isolated under Physiological Conditions. *Mol. Cell. Biol.* **26**: 5528-5543.

Domdey, H., Apostol, B., Lin, R., Newman, A., Brody, E., and Abelson, J. 1984. Lariat structures are in vivo intermediates in yeast pre-messenger RNA splicing. *Cell* **39**: 611-621.

Dougherty, W.G., Cary, S.M., and Parks, T.D. 1989. Molecular genetic analysis of a plant virus polyprotein cleavage site: a model. *Virology* **171**: 356-64.

Ducret, A., Van Oostveen, I., Eng, J.K., Yates, J.R., and Aebersold, R. 1998. High throughput protein characterization by automated reverse-phase chromatography/electrospray tandem mass spectrometry. *Protein science : a publication of the Protein Society* **7**: 706-19.

Enger, M.D., and Walters, R.A. 1970. Isolation of low molecular weight, methylated ribonucleic acids from 10S to 30S particles of Chinese hamster cell fractions. *Biochemistry* **9**: 3551-62.

Epstein, P., Reddy, R., Henning, D., and Busch, H. 1980. The nucleotide sequence of nuclear U6 (4.7 S) RNA. *The Journal of biological chemistry* **255**: 8901-6.

Fang, G., Rocha, E., and Danchin, A. 2005. How essential are nonessential genes? *Molecular biology and evolution* **22**: 2147-56.

Fantes, P., and Beggs, J. 2000. *The Yeast Nucleus*. 1st ed. Oxford University Press, USA.

Fenn, J.B., Mann, M., Meng, C.K., Wong, S.F., and Whitehouse, C.M. 1989. Electrospray ionization for mass spectrometry of large biomolecules. *Science (New York, N.Y.)* **246**: 64-71.

Fong, N., and Bentley, D. 2001. Capping, splicing, and 3 ' processing are independently stimulated by RNA polymerase II: different functions for different segments of the CTD. *Genes Dev.* **15**: 1783-1795.

Freund, C., Dötsch, V., Nishizawa, K., Reinherz, E.L., and Wagner, G. 1999. The GYF domain is a novel structural fold that is involved in lymphoid signaling through proline-rich sequences. *Nature structural biology* **6**: 656-60.

Galisson, F., and Legrain, P. 1993. The biochemical defects of prp4-1 and prp6-1 yeast splicing mutants reveal that the PRP6 protein is required for the accumulation of the [U4/U6.U5] tri-snRNP. *Nucleic acids research* **21**: 1555-62.

Gavin, A. et al. 2002. Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* **415**: 141-147.

Gietz, R., and Woods, R. 2002. Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. In *Guide to Yeast Genetics and Molecular and Cell Biology, Pt B*, *Methods in Enzymology*, pp. 87-96, 350.

Goffeau, A. et al. 1996. Life with 6000 Genes. *Science* **274**: 546-567.

Gonzalez-Santos, J.M., Wang, A., Jones, J., Ushida, C., Liu, J., and Hu, J. 2002. Central region of the human splicing factor Hprp3p interacts with Hprp4p. *The Journal of biological chemistry* **277**: 23764-72.

Görnemann, J., Kotovic, K.M., Hujer, K., and Neugebauer, K.M. 2005. Cotranscriptional spliceosome assembly occurs in a stepwise fashion and requires the cap binding complex. *Molecular cell* **19**: 53-63.

Gottschalk, A., Kastner, B., Luhrmann, R., and Fabrizio, P. 2001. The yeast U5 snRNP coisolated with the U1 snRNP has an unexpected protein composition and includes the splicing factor Aar2p. *RNA-Publ. RNA Soc.* **7**: 1554-1565.

Gottschalk, A., Neubauer, G., Banroques, J., Mann, M., Lührmann, R., and Fabrizio, P. 1999. Identification by mass spectrometry and functional analysis of novel proteins of the yeast [U4/U6•U5] tri-snRNP. *EMBO J.* **18**: 4535-4548.

Gottschalk, A., Bartels, C., Neubauer, G., Luhrmann, R., and Fabrizio, P. 2001. A Novel Yeast U2 snRNP Protein, Snu17p, Is Required for the First Catalytic Step of Splicing and for Progression of Spliceosome Assembly. *Mol. Cell. Biol.* **21**: 3037-3046.

Guthrie, C., and Patterson, B. 1988. Spliceosomal snRNAs. *Annual review of genetics* **22**: 387-419.

Harrison, P.M., Kumar, A., Lang, N., Snyder, M., and Gerstein, M. 2002. A question of size: the eukaryotic proteome and the problems in defining it. *Nucleic acids research* **30**: 1083-90.

Hausner, T., Giglio, L., and Weiner, A. 1990. Evidence for base-pairing between mammalian U2 and U6 small nuclear ribonucleoprotein particles. *Genes Dev.* **4**: 2146-2156.

Hodnett, J.L., and Busch, H. 1968. Isolation and characterization of uridylic acid-rich 7 S ribonucleic acid of rat liver nuclei. *The Journal of biological chemistry* **243**: 6334-42.

Hong, E.L. et al. 2008. Gene Ontology annotations at SGD: new data sources and annotation methods. *Nucleic acids research* **36**: D577-81.

Ito, T., Chiba, T., Ozawa, R., Yoshida, M., Hattori, M., and Sakaki, Y. 2001. A comprehensive two-hybrid analysis to explore the yeast protein interactome. *Proceedings of the National Academy of Sciences* **98**: 4569-4574.

Jády, B.E., and Kiss, T. 2000. Characterisation of the U83 and U84 small nucleolar RNAs: two novel 2'-O-ribose methylation guide RNAs that lack complementarities to ribosomal RNAs. *Nucleic acids research* **28**: 1348-54.

Jankowsky, E., Gross, C., Shuman, S., and Pyle, A. 2001. Active disruption of an RNA-protein interaction by a DExH/D RNA helicase. *Science* **291**: 121-125.

Jurica, M., and Moore, M. 2002. Capturing splicing complexes to study structure and mechanism. *Methods* **28**: 336-345.

Karas, M., and Hillenkamp, F. 1988. Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Analytical chemistry* **60**: 2299-301.

Kessler, S.W. 1975. Rapid Isolation of Antigens from Cells with A Staphylococcal Protein A-Antibody Adsorbent: Parameters of the Interaction of Antibody-Antigen Complexes with Protein A. *J Immunol* **115**: 1617-1624.

Kim, D.H., and Rossi, J.J. 1999. The first ATPase domain of the yeast 246-kDa protein is required for in vivo unwinding of the U4/U6 duplex. *RNA* **5**: 959-971.

Kim, E., Magen, A., and Ast, G. 2007. Different levels of alternative splicing among eukaryotes. *Nucl. Acids Res.* **35**: 125-131.

Kim, S., and Lin, R. 1993. Pre-mRNA splicing within an assembled yeast spliceosome requires an RNA-dependent ATPase and ATP hydrolysis. *Proc. Natl. Acad. Sci. U.S.A.* **90**: 888-892.

Kim, S., and Lin, R. 1996. Spliceosome activation by PRP2 ATPase prior to the first transesterification reaction of pre-mRNA splicing. *Mol. Cell. Biol.* **16**: 6810-6819.

King, D., and Beggs, J. 1990. Interactions of PRP2 protein with pre-mRNA splicing complexes in Saccharomyces cerevisiae. *Nucleic Acids Res.* **18**: 6559-64.

Kiss, A.M., Jády, B.E., Bertrand, E., and Kiss, T. 2004. Human box H/ACA pseudouridylation guide RNA machinery. *Molecular and cellular biology* **24**: 5797-807.

Kiss, T. 2004. Biogenesis of small nuclear RNPs. *J Cell Sci* **117**: 5949-5951.

Konarska, M., and Sharp, P. 1988. Association of U2, U4, U5, and U6 small nuclear ribonucleoproteins in a spliceosome-type complex in absence of precursor RNA. *Proc. Natl. Acad. Sci. U.S.A.* **85**: 5459-5462.

Konarska, M., and Sharp, P. 1987. Interactions between small nuclear ribonucleoprotein particles in formation of spliceosomes. *Cell* **49**: 763-774.

Konarska, M.M. 2008. A purified catalytically competent spliceosome. *Nat Struct Mol Biol* **15**: 222-224.

Kornberg, A. 1989. *For the Love of Enzymes: The Odyssey of a Biochemist*. Harvard University Press.

Krainer, A.R. 1997. *Eukaryotic mRNA Processing*. Oxford University Press, USA.

Krawczak, M., Reiss, J., and Cooper, D. 1992. The mutational spectrum of single base-pair substitutions in messenger RNA splice junctions of human genes - causes and consequences. *Hum. Genet.* **90**: 41-54.

Krol, A., Gallinaro, H., Lazar, E., Jacob, M., and Branlant, C. 1981. The nuclear 5S RNAs from chicken, rat and man. US RNAs are encoded by multiple genes. *Nucl. Acids Res.* **9**: 769-787.

Laggerbauer, B., Liu, S., Makarov, E., Vornlocher, H., Makarova, O., Ingelfinger, D., Achsel, T., and Luhrmann, R. 2005. The human U5 snRNP 52K protein (CD2BP2) interacts with U5-102K (hPrp6), a U4/U6.U5 tri-snRNP bridging protein, but dissociates upon tri-snRNP formation. *RNA* **11**.

Lander, E.S. et al. 2001. Initial sequencing and analysis of the human genome. *Nature* **409**: 860-921.

Lauber, J., Fabrizio, P., Teigelkamp, S., Lane, W., Hartmann, E., and Lührmann, R. 1996. The HeLa 200 kDa U5 snRNP-specific protein and its homologue in Saccharomyces cerevisiae are members of the DEXH-box protein family of putative RNA helicases. *EMBO J.* **15**: 4001-15.

Lerner, M.R., Boyle, J.A., Mount, S.M., Wolin, S.L., and Steitz, J.A. 1980. Are snRNPs involved in splicing? *Nature* **283**: 220-4.

Lerner, M., Boyle, J., Hardin, J., and Steitz, J. 1980. Two novel classes of small ribonuclearproteins detected by antibodies associated with lupus erythematosus. *Science* **211**: 400-402.

Lerner, M., and Steitz, J. 1979. Antibodies to small nuclear RNAs complexed with proteins are produced by patients with systemic lupus erythematosus. *Proc. Natl. Acad. Sci. U.S.A.* **76**: 5495-5499.

Libri, D., Graziani, N., Saguez, C., and Boulay, J. 2001. Multiple roles for the yeast SUB2/yUAP56 gene in splicing. *Genes Dev.* **15**: 36-41.

Lin, R., Newman, A., Cheng, S., and Abelson, J. 1985. Yeast mRNA splicing in vitro. *J Biol Chem* **260**: 14780-92.

Linder, P. 2006. Dead-box proteins: a family affair--active and passive players in RNP-remodeling. *Nucleic acids research* **34**: 4168-80.

Listerman, I., Sapra, A., and Neugebauer, K. 2006. Cotranscriptional coupling of splicing factor recruitment and precursor messenger RNA splicing in mammalian cells. *Nat. Struct. Mol. Biol.* **13**: 815-822.

Liu, S., Rauhut, R., Vornlocher, H., and Lührmann, R. 2006. The network of protein-protein interactions within the human U4/U6.U5 tri-snRNP. *RNA* **12**: 1418-1430.

Liu, S. 2005. Investigation of Protein-protein Interactions within the Human Spliceosomal U4/U6.U5 tri-snRNP Particle. Text.PhDThesis, Niedersächsische Staats- und Universitätsbibliothek Göttingen http://webdoc.sub.gwdg.de/diss/2005/liu/index.html (Accessed March 3, 2008).

Liu, Y., Chen, H., Wu, N., and Cheng, S. 2007. A Novel Splicing Factor, Yju2, Is Associated with NTC and Acts after Prp2 in Promoting the First Catalytic Reaction of Pre-mRNA Splicing . *Molecular and Cellular Biology* **27**. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1952081.

Liu, Z. 2002. p68 RNA Helicase Is an Essential Human Splicing Factor That Acts at the U1 snRNA-5' Splice Site Duplex. *Mol. Cell. Biol.* **22**: 5443-5450.

Luo, H.R., Moreau, G.A., Levin, N., and Moore, M.J. 1999. The human Prp8 protein is a component of both U2- and U12-dependent spliceosomes. *RNA* **5**.

Madhani, H., and Guthrie, C. 1992. A novel base-pairing interaction between U2 and U6 snRNAs suggests a mechanism for the catalytic activation of the spliceosome. *Cell* **71**: 803-17.

Madhani, H., and Guthrie, C. 1994. Dynamic RNA-RNA interactions in the spliceosome. *Annual Rev. Genet.* **28**: 1-26.

Makarov, E.M., Makarova, O.V., Achsel, T., and Luhrmann, R. 2000. The human homologue of the yeast splicing factor prp6p contains multiple TPR elements and is stably associated with the U5 snRNP via protein-protein interactions. *Journal of Molecular Biology* **298**: 567-575.

Makarova, O., Makarov, E., and Lührmann, R. 2001. The 65 and 110 kDa SR-related proteins of the U4/U6•U5 tri-snRNP are essential for the assembly of mature spliceosomes. *EMBO J.* **20**: 2553-2563.

Makarova, O., Makarov, E., Urlaub, H., Will, C., Gentzel, M., Wilm, M., and Lührmann, R. 2004. A subset of human 35S U5 proteins, including Prp19, function prior to catalytic step 1 of splicing. *EMBO J.* **23**: 2381-2391.

Makarova, O.V., Makarov, E.M., Liu, S., Vornlocher, H., and Lührmann, R. 2002. Protein 61K, encoded by a gene (PRPF31) linked to autosomal dominant retinitis pigmentosa, is required for U4/U6*U5 tri-snRNP formation and pre-mRNA splicing. *The EMBO journal* **21**: 1148-57.

Malca, H., Shomron, N., and Ast, G. 2003. The U1 snRNP base pairs with the 5' splice site within a penta-snRNP complex. *Molecular and cellular biology* **23**: 3442-55.

Mann, M., Hendrickson, R.C., and Pandey, A. 2001. Analysis of proteins and proteomes by mass spectrometry. *Annual review of biochemistry* **70**: 437-73.

Martin, A., Schneider, S., and Schwer, B. 2002. Prp43 is an essential RNA-dependent ATPase required for release of lariat-intron from the spliceosome. *J. Biol. Chem.* **277**: 17743-17750.

Massenet, S., Pellizzoni, L., Paushkin, S., Mattaj, I.W., and Dreyfuss, G. 2002. The SMN complex is associated with snRNPs throughout their cytoplasmic assembly pathway. *Molecular and cellular biology* **22**: 6533-41.

Matera, A.G., and Shpargel, K.B. 2006. Pumping RNA: nuclear bodybuilding along the RNP pipeline. *Current opinion in cell biology* **18**: 317-24.

Moore, M., Query, C., and Sharp, P. 1993. Splicing of Precursors to mRNA by the Spliceosome. In *The RNA World*, pp. 303-357, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N. Y.

Motoyama, A., Venable, J.D., Ruse, C.I., and Yates, J.R. 2006. Automated ultra-high-pressure multidimensional protein identification technology (UHP-MudPIT) for improved peptide identification of proteomic samples. *Analytical chemistry* **78**: 5109-18.

Muramatsu, M., Hodnett, J.L., and Busch, H. 1966. Base Composition of Fractions of Nuclear and Nucleolar Ribonucleic Acid Obtained by Sedimentation and Chromatography. *J. Biol. Chem.* **241**: 1544-1550.

Nielsen, T.K., Liu, S., Lührmann, R., and Ficner, R. 2007. Structural basis for the bifunctionality of the U5 snRNP 52K protein (CD2BP2). *Journal of molecular biology* **369**: 902-8.

Niggli, V., Penniston, J.T., and Carafoli, E. 1979. Purification of the (Ca2+-Mg2+)-ATPase from human erythrocyte membranes using a calmodulin affinity column. *The Journal of biological chemistry* **254**: 9955-8.

Nissim-Rafinia, M., and Kerem, B. 2005. The splicing machinery is a genetic modifier of disease severity. *Trends in genetics : TIG* **21**: 480-3.

Nottrott, S., Urlaub, H., and Lührmann, R. 2002. Hierarchical, clustered protein interact ions with U4/U6 snRNA: a biochemical role for U4/U6 proteins. *EMBO J.* **21**: 5527-5538.

van Nues, R.W., and Beggs, J.D. 2001. Functional Contacts With a Range of Splicing Proteins Suggest a Central Role for Brr2p in the Dynamic Control of the Order of Events in Spliceosomes of Saccharomyces cerevisiae. *Genetics* **157**: 1451-1467.

Ohno, M., Segref, A., Bachi, A., Wilm, M., and Mattaj, I.W. 2000. PHAX, a mediator of U snRNA nuclear export whose activity is regulated by phosphorylation. *Cell* **101**: 187-98.

Okamura, N., and Busch, H. 1965. Base Composition of High Molecular Weight Nuclear RNA of Walker Tumor and Liver of the Rat. *Cancer research* **25**: 693-7.

Padgett, R.A., Grabowski, P.J., Konarska, M.M., Seiler, S., and Sharp, P.A. 1986. Splicing of messenger RNA precursors. *Annual review of biochemistry* **55**: 1119-50.

Palacios, I., Hetzer, M., Adam, S.A., and Mattaj, I.W. 1997. Nuclear import of U snRNPs requires importin beta. *The EMBO Journal* **16**.

Parenteau, J. et al. 2008. Deletion of Many Yeast Introns Reveals a Minority of Genes that Require Splicing for Function. *Mol. Biol. Cell* ePub Ahead of Print.

Parker, R., Siliciano, P., and Guthrie, C. 1987. Recognition of the TACTAAC box during mRNA splicing in yeast involves base pairing to the U2-like snRNA. *Cell* **49**: 229-39.

Patel, A.A., and Steitz, J.A. 2003. Splicing double: insights from the second spliceosome. *Nature reviews. Molecular cell biology* **4**: 960-70.

Paushkin, S., Gubitz, A.K., Massenet, S., and Dreyfuss, G. 2002. The SMN complex, an assemblyosome of ribonucleoproteins. *Current opinion in cell biology* **14**: 305-12.

Popplewell, A.G., Gore, M.G., Scawen, M., and Atkinson, T. 1991. Synthesis and mutagenesis of an IgG-binding protein based upon protein A of Staphylococcus aureus. *Protein Eng.* **4**: 963-970.

Prestayko, A.W., Tonato, M., and Busch, H. 1970. Low molecular weight RNA associated with 28 s nucleolar RNA. *Journal of molecular biology* **47**: 505-15.

Puig, O., Caspary, F., Rigaut, G., Rutz, B., Bouveret, E., Bragado-Nilsson, E., Wilm, M., and Séraphin, B. 2001. The tandem affinity purification (TAP) method: A general procedure of protein complex purification. *Methods* **24**: 218-229.

Rader, S.D., and Guthrie, C. 2002. A conserved Lsm-interaction motif in Prp24 required for efficient U4/U6 di-snRNP formation. *RNA* **8**.

Raghunathan, P., and Guthrie, C. 1998a. A spliceosomal recycling factor that reanneals U4 and U6 small nuclear ribonucleoprotein particles. *Science* **279**: 857-860.

Raghunathan, P., and Guthrie, C. 1998b. RNA unwinding in U4/U6 snRNPs requires ATP hydrolysis and the DEIH-box splicing factor Brr2. *Current Biology* **8**: 847-855.

Raj, N.B., Ro-Choi, T.S., and Busch, H. 1975. Nuclear ribonucleoprotein complexes containing the U1 and U2 RNA. *Biochemistry* **14**: 4380-4385.

Reddy, R., Henning, D., and Busch, H. 1979. Nucleotide sequence of nucleolar U3B RNA. *The Journal of biological chemistry* **254**: 11097-105.

Reddy, R., Henning, D., and Busch, H. 1981. The primary nucleotide sequence of U4 RNA. *The Journal of biological chemistry* **256**: 3532-8.

Riedel, N., Wise, J.A., Swerdlow, H., Mak, A., and Guthrie, C. 1986. Small nuclear RNAs from Saccharomyces cerevisiae: unexpected diversity in abundance, size, and molecular complexity. *Proceedings of the National Academy of Sciences of the United States of America* **83**: 8097-101.

Rigaut, G., Shevchenko, A., Rutz, B., Wilm, M., Mann, M., and Séraphin, B. 1999. A generic protein purification method for protein complex characterization and proteome exploration. *Nature biotechnology* **17**: 1030-2.

Roy, J., Kim, K., Maddock, J.R., Anthony, J.G., and Jr, J.L.W. 1995. The final stages of spliceosome maturation require Spp2p that can interact with the DEAH box protein Prp2p and promote step 1 of splicing. *RNA* **1**.

Ruby, S., and Abelson, J. 1988. An early hierarchic role of U1 small nuclear ribonucleoprotein in spliceosome assembly. *Science* **242**: 1028-35.

Rutz, B., and Seraphin, B. 1999. Transient interaction of BBP/ScSF1 and Mud2 with the splicing machinery affects the kinetics of spliceosome assembly. *RNA-Publ. RNA Soc.* **5**: 819-31.

Schneider, B.L., Seufert, W., Steiner, B., Yang, Q.H., and Futcher, A.B. 1995. Use of polymerase chain reaction epitope tagging for protein tagging in Saccharomyces cerevisiae. *Yeast (Chichester, England)* **11**: 1265-74.

Schneider, C., Will, C.L., Makarova, O.V., Makarov, E.M., and Lührmann, R. 2002. Human U4/U6.U5 and U4atac/U6atac.U5 tri-snRNPs exhibit similar protein compositions. *Molecular and cellular biology* **22**: 3219-29.

Seraphin, B., Kretzner, L., and Rosbash, M. 1988. A U1 snRNA:pre-mRNA base pairing interaction is required early in yeast spliceosome assembly but does not uniquely define the 5' cleavage site. *Embo J* **7**: 2533-8.

Sharp, P.A. 2005. The discovery of split genes and RNA splicing. *Trends in Biochemical Sciences* **30**: 279-281.

Sherman, F. 1991. Getting Started with Yeast. *Methods in Enzymology* **194**: 3-21.

Shibata, H., Ro-Choi, T.S., Reddy, R., Choi, Y.C., Henning, D., and Busch, H. 1975. The primary nucleotide sequence of nuclear U-2 ribonucleic acid. The 5'-terminal portion of the molecule. *J. Biol. Chem.* **250**: 3909-3920.

Siliciano, P., and Guthrie, C. 1988. 5' splice site selection in yeast: genetic alterations in base-pairing with U1 reveal additional requirements. *Genes Dev.* **2**: 1258-1267.

Silverman, E., Maeda, A., Wei, J., Smith, P., Beggs, J., and Lin, R. 2004. Interaction between a G-patch protein and a spliceosomal DEXD/H-box ATPase that is critical for splicing. *Mol. Cell. Biol.* **24**: 10101-10110.

Small, E.C., Leggett, S.R., Winans, A.A., and Staley, J.P. 2006. The EF-G-like GTPase Snu114p regulates spliceosome dynamics mediated by Brr2p, a DExD/H box ATPase. *Molecular cell* **23**: 389-99.

Sontheimer, E.J. 2001. The spliceosome shows its metal. *Nat Struct Mol Biol* **8**: 11-13.

Staley, J.P., and Guthrie, C. 1999. An RNA switch at the 5' splice site requires ATP and the DEAD box protein Prp28p. *Molecular cell* **3**: 55-64.

Staley, J., and Guthrie, C. 1998. Mechanical devices of the spliceosome: Motors, clocks, springs, and things. *Cell* **92**: 315-326.

Stevens, S.W., Barta, I., Ge, H.Y., Moore, R.E., Young, M.K., Lee, T.D., and Abelson, J. 2001. Biochemical and genetic analyses of the U5, U6, and U4/U6 x U5 small nuclear ribonucleoproteins from Saccharomyces cerevisiae. *RNA* **7**: 1543-1553.

Stevens, S., and Abelson, J. 2002. Yeast pre-mRNA splicing: Methods, mechanisms, and machinery. In *Methods Enzymol.*, *Methods Enzymol.*, pp. 200-220, 351.

Stevens, S.W., and Abelson, J. 1999. Purification of the yeast U4/U6.U5 small nuclear ribonucleoprotein particle and identification of its proteins. *Proceedings of the National Academy of Sciences* **96**: 7226-7231.

Stevens, S.W., Ryan, D.E., Ge, H.Y., Moore, R.E., Young, M.K., Lee, T.D., and Abelson, J. 2002. Composition and Functional Characterization of the Yeast Spliceosomal Penta-snRNP. *Molecular Cell* **9**: 31-44.

Stirling, D.A., Petrie, A., Pulford, D.J., Paterson, D.T.W., and Stark, M.J.R. 1992. Protein A-calmodulin fusions: a novel approach for investigating calmodulin function in yeast. *Molecular Microbiology* **6**: 703-713.

Tanner, N., and Linder, P. 2001. DExD/H box RNA helicases: From generic motors to specific dissociation functions. *Molecular Cell* **8**: 251-262.

Tardiff, D., and Rosbash, M. 2006. Arrested yeast splicing complexes indicate stepwise snRNP recruitment during in vivo spliceosome assembly. *RNA* **12**: 968-979.

Tarn, W., Hsu, C., Huang, K., Chen, H., Kao, H., Lee, K., and Cheng, S. 1994. Functional association of essential splicing factors with Prp19 in a protein complex. *EMBO J.* **13**: 2421-2431.

Tarn, W., Lee, K., and Cheng, S. 1993. The yeast Prp19 protein is not tightly associated with small nuclear RNAs, but appears to associate with the spliceosome after binding of U2 to the premessenger RNA and prior to formation of the functional spliceosome. *Mol. Cell. Biol.* **13**: 1883-1891.

Tarn, W., and Steitz, J.A. 1997. Pre-mRNA splicing: the discovery of a new spliceosome doubles the challenge. *Trends in Biochemical Sciences* **22**: 132-137.

Thomas, K. 2007. Beer, and the Biochemists Behind it. *The Scientist*. http://www.the-scientist.com/news/display/52923/.

Thompson, J. 1907. On Rays of Positive Electricity. *Phil.Mag.S.* **6**: 561-575.

Treisman, R., Orkin, S.H., and Maniatis, T. 1983. Specific transcription and RNA splicing defects in five cloned beta-thalassaemia genes. *Nature* **302**: 591-6.

Tsai, W. et al. 1999. Cef1p is a component of the Prp19p-associated complex and essential for pre-mRNA splicing. *J. Biol. Chem.* **274**: 9455-9462.

Uetz, P. et al. 2000. A comprehensive analysis of protein-protein interactions in Saccharomyces cerevisiae. *Nature* **403**: 623-627.

Umen, J., and Guthrie, C. 1995. The second catalytic step of pre-mRNA splicing. *RNA* **1**: 869-85.

Villa, T., Pleiss, J., and Guthrie, C. 2002. Spliceosomal snRNAs: Mg2+-dependent chemistry at the catalytic core? *Cell* **109**: 149-152.

Vincent, K., Wang, Q., Jay, S., Hobbs, K., and Rymond, B. 2003. Genetic interactions with CLF1 identify additional pre-mRNA splicing factors and a link between activators of yeast vesicular transport and splicing. *Genetics* **164**: 895-907.

Wagner, J., Jankowsky, E., Company, M., Pyle, A., and Abelson, J. 1998. The DEAH-box protein PRP22 is an ATPase that mediates ATP- dependent mRNA release from the spliceosome and unwinds RNA duplexes. *EMBO J.* **17**: 2926-2937.

Wang, Q., and Rymond, B.C. 2003. Rds3p is required for stable U2 snRNP recruitment to the splicing apparatus. *Molecular and cellular biology* **23**: 7339-49.

Washburn, M., Wolters, D., and Yates, J. 2001. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nature Biotechnology* **19**: 242-247.

Wassarman, D., and Steitz, J. 1993. A base-pairing interaction between U2 and U6 small nuclear RNAs occurs in >150s complexes in HeLa-cell extracts - implications for the spliceosome assembly pathway. *Proc. Natl. Acad. Sci., U.S.A.* **90**: 7139-7143.

Wassarman, K., and Steitz, J. 1993. Association with terminal exons in pre-messenger RNAs - a new role for the U1 snRNP. *Genes Dev.* **7**: 647-659.

West, R.W., and Milgrom, E. 2002. DEAD-box RNA helicase Sub2 is required for expression of lacZ fusions in Saccharomyces cerevisiae and is a dosage-dependent suppressor of RLR1 (THO2). *Gene* **288**: 19-27.

Wu, J., and Manley, J. 1989. Mammalian pre-mRNA branch site selection by U2 snRNP involves base pairing. *Genes Dev.* **3**: 1553-1561.

Xu, D., Nouraini, S., Field, D., Tang, S., and Friesen, J. 1996. An RNA-dependent ATPase associated with U2/U6 snRNAs in pre-mRNA splicing. *Nature* **381**: 709-13.

Yamashita, M., and Finn, J. 1984. Electrospray ion source. Another variation on the free-jet theme. *The Journal of Physical Chemistry* **88**: 4451-4459.

Yean, S.L., Wuenschell, G., Termini, J., and Lin, R.J. 2000. Metal-ion coordination by U6 small nuclear RNA contributes to catalysis in the spliceosome. *Nature* **408**: 881-4.

Yean, S., and Lin, R. 1991. U4 small nuclear RNA dissociates from a yeast spliceosome and does not participate in the subsequent splicing reaction. *Mol. Cell. Biol.* **11**: 5571-7.

Yoshinari, S., Itoh, T., Hallam, S.J., DeLong, E.F., Yokobori, S., Yamagishi, A., Oshima, T., Kita, K., and Watanabe, Y. 2006. Archaeal pre-mRNA splicing: a connection to hetero-oligomeric splicing endonuclease. *Biochemical and biophysical research communications* **346**: 1024-32.

Zhou, Z., and Reed, R. 1998. Human homologs of yeast prp16 and prp17 reveal conservation of the mechanism for catalytic step II of pre-mRNA splicing. *Embo J* **17**: 2095-106.

Zieve, G., and Penman, S. 1976. Small RNA species of the HeLa cell: metabolism and subcellular localization. *Cell* **8**: 19-31.

VITA

Andrew Adam Roth earned a B.S. from the University of Tennessee in Engineering Science and Mechanics concentrating in Biomedical Engineering. After working in the Vascular Research Laboratory at the University of Tennessee Medical Center, he reentered school at the University of Colorado. In 2000 he earned an MBA and was subsequently employed as a Financial Analyst with the Intel Corporation. He entered the graduate school at the University of Texas as a student with the Institute for Cellular and Molecular Biology in 2002.

Permanent address: 512 Rocky River Rd, Austin TX 78746

This dissertation was typed by the author.